WHAT IS USER FRIENDLY?

WHAT IS USER FRIENDLY?

Papers presented at the 1986 Clinic on Library Applications of Data Processing, April 20-22, 1986

Clinic on Library Applications of Data Processing: 1986

What is User Friendly?

Edited by
F.W. LANCASTER

Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign

ISBN 0-87845-072-6 ISSN 0069-4789

CONTENTS

Introduction

Considerable emphasis is now being given to the design or redesign of online systems in order to make them "easier to use" and thus "more attractive" to potential users. But do we really know what is meant by "easier to use" and "more attractive" in this context? These were the questions addressed at the Twenty-Third Annual Clinic on Library Applications of Data Processing, held in the Levis Faculty Center, University of Illinois at Urbana-Champaign, on April 20-22, 1986.

This volume contains the texts of the papers presented at the meeting. All of the authors explore the idea of "user friendly" as it applies to online catalogs and related tools. Some of them summarize their own experiences in the implementation of online systems in academic and public libraries, some look at the broader psychological and social aspects of interaction between users and systems, and some attempt to predict what the future may hold for online bibliographic systems.

This general overview of user friendly interface design should be of interest to all managers, systems analysts, consultants and other professionals involved in the planning, development, and use of automated systems in libraries and information centers of all types.

F.W. LANCASTER
Editor

MICHAEL GORMAN

Director of General Services and
Professor of Library Administration
University of Illinois at Urbana-Champaign

Linking the Unlinkable

The best advice I know on after-dinner speaking comes from a book of the Apocrypha:

> Let thy speech be short, comprehending much in a few words (Ecclesiasticus, 22:8).

It was almost exactly ten years ago when I first spoke at a Clinic on Data Processing. At that time, I was merely a humble spear-carrier—a mere paper-giver. The keynote speech on that occasion was given by Frederick Kilgour. It was the first time that I had seen that eminent gentleman. He looked, then as now, more like the senior senator from Ohio than one of the leading innovators of modern librarianship. The years have rolled by and I find myself with the daunting task of following in the distinguished footsteps of the likes of Mr. Kilgour. Though my hair lacks the true senatorial silveriness which so distinguishes Fred, it has much more gray than it had in 1976. The amount of that gray which is not due to heredity is due, in large part, to wrestling with the principles and practicalities of the online catalog (as we call it for convenience). It is the implications of one aspect of that automated bibliographic control system that I wish to discuss this evening. Specifically, the burden of my song is the idea of using microcomputers as the central component of a third way of achieving and extending developed online catalogs. (Incidentally, I must take full responsibility for the title of this keynote speech. My fondness for facetious titles [and, indeed, for facetiousness in general] has not dimmed with the years and I forced my waggishness upon my distinguished compatriot Professor Wilfrid Lancaster, who is hereby absolved of all responsibility.)

I referred, a little while ago, to the fact that the term *online catalog* is now simply a term of convenience and one which is now so inaccurate as to be seriously misleading. (By the way, though I find the term unsatisfactory, I still prefer it to the horrid acronym OPAC [for "online public access catalog"]. Quite apart from the overtones of OPEC, there is the idea that an OPAC might be a political action committee which is dedicated to nothing—the zero-PAC.) The idea has never been that we should simply automate the pre-machine catalog (though, to tell the truth, some have tried to do just that), but that we should produce an online system which has at least three important differences from the pre-machine catalog.

The first of these major differences is that the online system should be more responsive to the needs of the library user than is, say, the card catalog and will allow many more ways of obtaining the information which is held in the system. This is readily achievable since even the worst computer systems are more responsive and forgiving than the card catalog ever was. Second, the online system should be more available to the user than its predecessors. By and large, we have achieved this second aim too by siting terminals in various locations in our libraries and communities. This has not been an invariable practice. Some libraries have been influenced by some rather rum "studies" of catalog use which have demonstrated conclusively that library users use card catalogs in places where those catalogs are sited. This clearly proves, to some, that terminals should be situated in the same place as the card catalog. This zany logic leads to a loss of one of the great advantages of the online system. The third important difference, and the one with which I am primarily concerned this evening, is that the online system will contain far more information than its predecessors.

In order to understand and examine this last point, we need to look at the situation which the users faced in using pre-machine systems. The fundamental problem was that the user's expectations were far higher than the capabilities of the bibliographic control system. He or she expected to be able to use the catalog to determine the availability of the materials sought. The catalog was not concerned with questions of availability but with questions of ownership. The user's question is, "*Ubi est meum*?" ("Where is mine?"—Mike Royko's proposed motto for the city of Chicago). The pre-machine catalog's dusty answer was "The library owns, or believes it owns, this item." It has been amply demonstrated, in libraries and in the wider world, that, when answers do not match questions, a crisis of confidence results. The well-kept secret was, of course, that the information which was needed to answer the users' questions was scattered throughout numerous other files created and maintained by the library. The on-order file, the binding file, the circulation file, the serial record, the serial check-in file...the list of these public and private files was as extensive as it was dreary. Few librarians knew the ins and outs of all of these and almost all users were blissfully ignorant of their very existence. When I first came to the University of Illinois Library, my then-assistant did a census of the paper files maintained in the technical services departments. They were more than sixty in number, of varying sizes and purposes. My favorite was the "Dead Slavic Serials File." Surely the only thing on God's earth which is sadder than a dead Slavic serial is the memorial within which its demise is recorded! The task of the online replacement for the pre-machine catalog is to bring all this scattered information together and to make it available to the library user.

Since the beginning of computerized bibliographic systems in libraries there has been a perception that there are two ways of bringing all this previously scattered information together. To simplify, the discussion has centered on the choice between integrated and separate systems with the smart money tending to favor the first. By an integrated system is meant one in which all the information about the materials held by or ordered by the library is stored and manipulated by an integrated set of programs within a single hardware configuration. Further, in an integrated system all this information is presented to the user at one terminal. On the other hand, separate systems would be those in which each function is carried out independently of each other function. Some of these separate systems may even require separate central computers and separate and different terminals. This is an over-simplified picture because what has happened often in the real world has been that many libraries have created a hybrid of partially integrated and partially separate systems. In this latter case, for example, the functions of the catalog and the circulation system might be integrated and the acquisitions and serial control functions might each be carried out by separate systems.

Although the integrated system has been seen by most as the preferred alternative, the fact remains that few if any truly integrated systems have been achieved in medium-sized or large libraries. Even the partially integrated systems that have been achieved have been bedeviled by the complexity of the software which is required to deal with a number of interrelated subsystems. Fitting the different data required for different functions into the Procrustean bed of the integrated system format has proved to be even more difficult. The concept of separate systems for separate functions has not been favored because it makes more work for the library user and because it is really no more than the automated version of the pre-machine systems. On the other hand, there are distinct advantages for the programmer and system specifier when it comes to creating a tailor-made system to carry out a specific function. The choice, then, has always seemed to be between the complex architecture of the integrated system and the user hostility of the separate system approach. There is, however, another possibility which may resolve the seemingly inescapable dilemma. That third way is made possible by the use of the microcomputer.

There is another dimension to this matter. It concerns the need for information other than the traditional bibliographic information found in catalogs, order files, and the like. That information consists of information about serial articles (from indexes, abstracts, etc.), data in electronic form, and (though it is little more than a gleam in Fred Kilgour's eye) the full text of publications in electronic form. It is hard even to imagine the integrated system which would bring all this and the traditional kinds of bibliographic information together and even harder to realize such a system. It is almost depressing to think of the separate system concept applied in this area. The thought of the library user being presented with twenty or more different terminals, each with its own commands and demands, is dismal indeed. Such an electronic Maginot Line would require staff resources which few libraries possess and would demand more application and effort from the user than any library has a right to expect. I wish I were as modern and progressive in these matters as Wilf Lancaster, but the fact remains that I still cling to the idea of the library more or less as we know it, to the notion that library service is intimately connected to the provision of information about printed materials (books, serials, etc.) as well as to the more whiz-bang materials, and to the belief that new methods of communication supplement rather than replace the older forms.

The germ of the Third Way—the alternative to both the integrated and separate system concepts—was born of the dilemma which we faced at the UIUC Library in combining a circulation system (LCS) and a MARC-based bibliographic system (WLN—Washington Library Network) to form our online catalog. For the moment all I need to mention is that we rejected both the idea of integrating the two systems (in any event a perilous and uncertain venture) and the idea of maintaining both as separate systems (if for no other reason than this approach would have been unfriendly, to say the least, to the library user). What we have done is to use microcomputers (IBM PCs) as public terminals, to implant interface programs in those microcomputers which translate the user's natural language queries into the arcane commands of the two systems, and to set up interactions between the microcomputers and the mainframe computer which economize on telecommunication costs (in that the majority of the processing—and all the unproductive processing—is done in the microcomputer). This is a small step for one library but one that is not without significance for library kind. The significance lies not in our local application but in the fact that two quite different systems are presented to the user as if they are one system. They have not been integrated but they do not stand alone. The circle has been squared. Neither integrated nor separate, the systems are nevertheless in harmony with the needs of the user. Remember also that these are completely different systems each with its own deep structure and each with its own economy and purpose.

I would suggest that this modest beginning opens up important possibilities for all online bibliographic systems and for the provision of the kind of nonbibliographic information which I mentioned earlier. The essential point is that if, as we have demonstrated, one can design and write an interface program which links two completely different bibliographic systems then one could write such programs to link three, four, or five, or more such systems. In other words, the advantages of the separate system (that it is tailor-made for a particular function and performs its tasks with economy and efficiency) can be maintained in an environment which presents the user with the advantages of the integrated system (the bringing together and display of hitherto scattered and secret information).

Having thus resolved the dilemma of integrated *v.* separate bibliographic systems, let us turn our attention to the nonbibliographic dimension. This comprises three classes. The first is that of serial literature (what Dr. Ranganathan called "microthought"). We have traditionally given access to this kind of publication by means of printed indexing and abstracting services and latterly by online versions of such services. These services are inefficient, to say the least, because they are unorganized, random to a great degree, and because they are completely separate from the traditional bibliographic systems of the library. This is caused, in great part, by the fact that the indexing and abstracting services emanate from the for-profit sector. That sector is almost always philosophically and practically out of tune with the nonprofit sector which includes most libraries. The microcomputer, used intelligently, offers a way out of this problem too. If one can use a microcomputer to interact between two or more incompatible bibliographic systems, then there is no reason why its use could not be extended to the interaction between bibliographic and indexing/abstracting systems. Those services could be either online or held as a local database using videodisc technology.

Such an interaction of systems would go a long way, I believe, to refuting certain anomalous and erroneous findings of studies of early online catalogs. Those findings indicate that subject heading use increases dramatically when the move is made from pre-machine to online catalogs. It is my firm belief that this is a transitory phenomenon and that the increase in subject searches is partly due to the novelty of the online catalog and, in great part, to the fact that nothing better is available. I would predict confidently that, given easy and free subject access to current serial literature online (as part of the microcomputer-coordinated total library system), subject searches for monograph literature would subside to the previous low level. The key words in the preceding sentence are "easy and free subject access." The question of making the access easy for the user (to conform to Mooers Law of Least Effort) is technical and relatively easily solved. The question of free access is one which is financial, strategic, and political. It involves the reconciliation of the for-profit and nonprofit sectors and can thus be regarded as, at very least, thorny. On the other hand, if we are serious about using technology to move into a new dimension of library service, then I can see no better struggle upon which to embark.

The second nonbibliographic class is that of data itself. There is, as has been pointed out often, an ever-growing mass of data available in machine-readable form. This data is not only available but is also, given the right programs, manipulable by the user. Again, there is no technical reason why such data and such programs could not be made available to the user, at the same terminal as the bibliographic and serial information, by the microcomputer controlled library system. This availability could be secured either to databases at remote locations or to locally held databases (again, perhaps using videodisc technology).

Lastly, there is the question of the electronic publication (monographic and serial in nature). Fred Kilgour (whose benign presence pervades this paper) is currently engaged upon a research project called EIDOS which seeks to make the content of monographs in machine-readable form available to the user. This access will be primarily by "unconventional" means (searches of contents pages, captions, full text, etc.). Such techniques could be applied, together with more conventional access points, to serial publications in machine-readable form. When EIDOS is operational and when the volume and importance of electronic journals merits it, the microcomputer-controlled library system will reach out to engage these sources of information and knowledge and to bring them to the user.

My message, then, is that the process of integrating and bringing to light the hitherto scattered information about library materials is most successfully achieved by microcomputer coordination of separate and differing systems rather than by attempts at completely integrated library systems. Beyond this, that the quantum leap in service which has been the result of the creation of "online catalogs" will be matched and exceeded by the next generation of library systems. Those systems will not only deal with bibliographic information but will also embrace the worlds of microthought, of data, and of publications in machine-readable form. All of this adds up neither to the demise of the library nor to the replacement of traditional means of communicating information and knowledge. On the contrary, it will lead to hitherto undreamed-of levels of enhancement of library service. Many years ago, Charles Ammi Cutter lamented the end of "the golden age of cataloguing." It is my firm belief that the library is on the threshold of a new Golden Age of bibliographic control and of provision of nonbibliographic information, and that a prime tool in this renaissance will be the humble microcomputer.

Post scriptum: Since this paper was delivered on a Sunday and since it opens with a quotation from a book of the Jerusalem Bible, it seems fitting to record a Biblical quotation (for which I am indebted to Lowell Oxtoby of Western Illinois University Library) on the topic of the importance of redundancy in computer systems:

> Two are better than one; because they have a good reward for their labour.
>
> For if they fall, the one will lift up his fellow: but woe to him that is alone when he falleth; for he hath not another to help him up (Ecclesiastes, 4:9-10).

The autonomy and importance of the microcomputer in the systems which I envisage makes this exhortation of peculiar relevance.

WARD SHAW
Director
Colorado Alliance of Research Libraries

Aristotle Meets Plato in the Library Catalog: Part 1

This paper is part 1 of a presentation titled "Aristotle Meets Plato in the Library Catalog." In it, I hope to set forth some aspects of the theoretical context, or point of view, from which we at the Colorado Alliance of Research Libraries (CARL) approach the design and implementation of what the organizers of this clinic have called "user friendly" systems, to describe a bit the organizational and systems setting within which we work, outline some of the design principles that guide our development, and provide a brief overview of the system as it exists today. Part 2, by Ken Dowlin, will discuss the system in an application context at the Pikes Peak Library District in Colorado Springs. The system in question is one developed by the Colorado Alliance of Research Libraries, and available for installation elsewhere through Eyring Research Institute, to whom we have granted a marketing license. It forms the basis for MAGGIE III, the system in Colorado Springs.

The Theoretical Context

You will be relieved to learn that this is not Philosophy 101. However, as we try to address the question "What is user friendly?" it is important to uncover some basic assumptions that underlie our particular implementation of a public system. Let us examine our theoretical context, with the understanding that all of it is emphatically arguable. First, a public catalog is an information system. Information is the name of a process; specifically, the process by which people become informed. The process by which people become informed is closely related to, or maybe the same as, learning. The name for sparking learning is *teaching*. Hence, an important characteristic of public systems is that they teach, and one measure of their utility is the effectiveness with which they teach. Teaching, as any who follow debates surrounding educational policy will appreciate, is not well understood. One is led inevitably to the conclusion that we do not know what we are designing or at least that we do not have any guaranteed rules to follow.

Aristotle was the champion of the *a posteriori* method. If he wished to learn about a triangle, for example, he would analyze its parts and the mechanisms of their assembly by observation. He invented classification and, for all but the name, the scientific method. Plato, on the other hand, concentrated on the *a priori* method of learning. If he wished to learn about a triangle, he would consider its "triangleness" and draw logical conclusions from that concept. For him, the whole was both greater than and different from the sum of its parts.

We have applied Aristotelean methodology with considerable skill and marvelous detail in the construction of our classification schemes, MARC records, analytics, authorities, etc. in the design of research libraries and their traditional access tools. The method has served us remarkably well in providing conceptual structures for managing and controlling enormous resources, and its use was dictated by the technologies available. The difficulty, of course, is that the tools we have constructed are complex in direct relation to the fineness of the analysis they represent and require of their users intimate knowledge of system structure as well as discipline structure. Divergence of the two structures is inevitable and extremely difficult to control. Part but not all of this difficulty is, to be sure, a function of the relatively inflexible (expensive) technology of their traditional implementation. A large portion of the problem is that one must force one's thinking into the analytic patterns upon which the system is constructed, and it is thus exceedingly difficult to have new ideas. As McCluhan says, "the medium resists" and mightily.

Martin Heidegger, a twentieth-century phenomenologist, has written and spoken in detail about the concept of a tool, pointing out that a hammer, in the hands of a carpenter, is an extension of his arm. The carpenter uses the hammer to drive nails with wonderful efficiency and without thinking about it. While I do not have to know much about the hammer to pick it up, I must think about it in detail before I use it if only to avoid pounding my thumb. But Heidegger says that I am much more likely to conceive new uses for the hammer precisely because I see it as a tool for pounding rather than as a tool for driving nails. I look at its "hammerness," as might Plato, and draw *a priori* conclusions from the concept. In this case the medium also resists but potentially productively. And the more general (or "platonic") the tool, the more productive the resistance might be.

With electronic technology, the challenge is to enable users to manipulate our Aristotelian structures in Platonic forms, driving the systems to explore what users conceive rather than what we have "analyzed in." That we believe is the heart of "user friendliness" and is the sense in which we offer the title of this presentation. It is the basic context from which we attempt design.

The Organizational Context

CARL is a private nonprofit corporation in Colorado that has as members the libraries of the University of Colorado at Boulder, the University of Northern Colorado, the University of Denver, the Colorado School of Mines, as well as the Denver Public Library, and the Auraria Library which itself serves a consortium of three institutions of higher education in Denver. These are different kinds of organizations. They are state-supported, city-supported, and private. They are large general academic, large public, small special academic libraries. They differ in size from the University of Colorado—Boulder Library, a member of ARL; to the School of Mines Library serving a specialized academic clientele. They are also alike in certain important ways. They all have, as a part of their reason for being, the need to support graduate-level research, they all support large numbers of undergraduate students, and they all have a commitment of one kind or another to serve a wider user population than that of their immediate campus or city.

Governance of CARL is via its Council of Members consisting of the directors of each of the member libraries. In addition, CARL has a board of directors (not to be confused with the library directors), but in practice policy is set by the council.

CARL exists to create a single research resource for the various publics served by the member institutions. Said another way, CARL manages the collections of member institutions as if they were one collection. In order to accomplish this we have undertaken a whole series of network programs. The Colorado Organization for Library Acquisitions (COLA), for example, is a CARL program for cooperative acquisition of expensive material. It differs somewhat from other similar efforts in that the material purchased, although housed in the member libraries, is owned by CARL. We are developing a considerable collection enhancing those of the members. We also cooperatively purchase supplies and equipment for the members, when volume can generate savings. CARL's major program is the network online system. In order to create a single research resource, we needed one common mechanism to identify, locate, and control items throughout the network, and we also needed (and still need) a system for rapid, site-independent document delivery.

Design Principles

In creating our system, we attend to several design principles which are derivable from the theoretical and organizational contexts just described. I haven't time, obviously, to discuss them in detail but will briefly outline a few of them to provide a flavor of our approach.

First, the approach we use is heuristic, rejecting the algorithmic and simulation approaches as variously cumbersome, slow, and requiring impossible degrees of prior specification. As a result, our design principles are essentially statements of supposed value, and in some cases they are in direct conflict with each other. Each should be preceded by some substantial qualifier such as "generally, in most cases, it is probably the case that...." Negotiating between the principles requires constant trade-offs and modifications. Some principles regarding the overall system follow:

—The system must make it easy for users to view the network as a whole.
—The system must support local differences in both policy and practice.
—The system must promote experimentation.
—The system must provide very fast response time.
—The system ought not to require the user to understand the structure of a bibliographic record or of its associated files but rather ought to promote and support the construction of his own concept of organization. (We are indebted to Christine Borgman for alerting us to the idea of the user's "conceptual map.")
—The user must feel in control of the system, and not the other way around.
—The system must adapt to the user's skill level.
—The user should be able to get real results very quickly and then be able to experiment with variations very easily so that he may use the system to "explore."

Some principles from a hardware/software point of view are:

—Both hardware and software must be modular in design, allowing relatively easy changes to part of the system without dire consequences for the rest.
—Pass constant values to software as data.
—Separate message content and message form.
—Keep data structures flexible.
—Minimize disc accesses.

And some principles relating to screen design:

—Avoid library jargon and especially avoid computer jargon.
—Keep screens uncluttered.
—Avoid cuteness.

—Provide a cursor at the spot user typing will appear, and make that spot consistent from screen to screen.
—Don't tell users they have done something wrong. Rather, let results speak for themselves and provide positive suggestions. Assume that users are in control.
—Don't use blinking fields or reverse video.
—Systems have style. Keep it consistent.
—Pay attention to layout as well as content.

In summary, users know best what they do albeit sometimes with considerable professional help. System designers know best what the system can do. The goal of user friendliness is to provide a powerful, flexible, informative way for users to drive and control the system to their various ends. It is emphatically not to presume their ends or to channel their thinking according to predefined routes.

The System Overview

I'd like now to give you an overview of the CARL installation and the software (Pikes Peak Library District has a different configuration). The CARL hardware base is an eight-processor Tandem Nonstop II system. Each processor has 4 million bytes of main memory—32 million for the current system. There are 6 billion bytes of disc memory for the files. In the six library sites, 390 terminals communicate with the system via various network communications equipment.

Bibliographic records in the database come from all six institutions. From the system point of view, these records are organized in a common way and each field in each record contains an ownership bit map to indicate which institution "owns" which field. From the user's point of view, however, the records are organized by institution—that is, the user searches and examines records one institution at a time. Early versions of the system required a cumbersome reentry of each search when switching from one institution's files to another's, and more recently we have made that switching extremely easy. Ultimately we will support global searches. This progression was designed for political reasons. Individual institutions are wary of potential work loads on less heavily worked library subsystems such as interlibrary loans created by users from other institutions looking directly at their records. This fear has eased considerably with experience, partly because users who identify items they want at other institutions tend to go there directly rather than use traditional interlibrary methods to get the material. As these perceptions have changed, the system has changed to reflect (lead?) new concepts.

The software is organized into four distinct modules. First, the background software builds the database and creates the necessary indexing. Records are taken from OCLC, Autographics, and one or two other sources that members create as a result of their own cataloging activities. The software converts these records into our internal format and maintains the appropriate indexes. The various local fields are processed to create item records for circulation. The second software module is the public access catalog or PAC which provides searching of and switching between whatever data are resident on the system.

The third program module is the circulation system. This is a full-service system supporting charge, return, inquiry, holds, recall, tracers, overdues, fines, lists, letters, reserves, conversion, statistical reports, and secured full edit control over all files and records. We interface directly with various academic computing centers for the transmission of accounting data generated by system activity. Of primary importance is that circulation status of items shows up instantly in PAC so that users have up-to-the-minute information about availability of items they discover.

The fourth software module is bibliographic maintenance. Maint, as we call it, is used primarily for editorial changes to the MARC records. All fields are fully editable, and the program performs format checking and correction where appropriate to ensure MARC compatibility. Additionally, users can add and delete records. All changes are immediately processed and reflected in PAC and CIRC.

The fifth module is acquisitions. Currently ready for beta test in one of the member libraries, it is scheduled for systemwide installation in the summer. The sixth module, serials control, is now in design. User access to these modules, as well as to Tandem or locally developed services, is available and secured through NEWPEX.

The CARL database at the moment contains 1.85 million institution-unique bibliographic records and perhaps 3 million holdings. In addition to the 300 dedicated terminals, we provide free dial-up access to PAC, currently handling about 150 calls per day. We average about 1,200,000 message transactions per day with an average response time of .4 seconds. By the end of 1986, we anticipate a database of 2.5 million records and 450 dedicated terminals, generating 1.8 million daily transactions. Over 20,000 people use the system on a typical day.

KEN DOWLIN
Director
Pikes Peak Library District

Aristotle Meets Plato in the Library Catalog: Part 2

I will discuss the Colorado Alliance of Research Libraries (CARL) system implementation as the basic housekeeping system at the Pikes Peak Library District (PPLD) on MAGGIE III and how PPLD has used the capabilities of the software in the CARL system to greatly expand and enhance MAGGIE'S PLACE (the computer system at PPLD). But first, a recap of the history of MAGGIE'S PLACE.

MAGGIE'S PLACE

The automation program at PPLD was started in 1975 and the first in-house computer, a Digital Equipment Corporation (DEC) PDP 11/70, was acquired in 1976. This computer was dubbed MAGGIE II with its namesake the long-time head of the technical services department, Margaret O'Rourke. Over the next five years, through the hard work of the employees of the systems division of the library, MAGGIE'S PLACE became one of the most comprehensive and sophisticated library automated systems in existence. Program implementation started with a collection inventory system, then proceeded through circulation, acquisitions, serials, a public access catalog, and continued on to payroll, accounting, word processing, electronic mail, and other housekeeping tasks. PPLD broke new ground for the library world when community resource files were brought up on MAGGIE II in 1978. These files contain community agencies, clubs and organizations, adult education courses, an events calendar, and day-care centers. In 1981 PPLD added a second DEC computer, a PDP 11/44, to initiate the first community-wide public online CARPOOL system in the United States and later brought up a transit information system which provides the online schedules of the city bus system. Also in 1981 PPLD became the first library to allow owners of home microcomputers and business computers to link with MAGGIE for searching the public access catalog and the community resource files. By 1983 the catalogs at the libraries of the U.S. Air Force Academy and the University of Colorado at Colorado Springs were online in addition to the eight facilities of PPLD.

By 1983 MAGGIE had grown into a system with ninety-seven terminals and 1.8 billion bytes of storage. Unfortunately, this was above the maximum effective capacity for an 11/70 in the kind of activities required by PPLD. In 1983 the voters of PPLD approved a bond issue for a facility that would increase the total square footage of the district by 80 percent. This megabranch would require a minimum of fifty terminals—an impossible addition to MAGGIE II. Fortunately the bond issue included funds for an entirely new computer system. After the passage of the bond issue the specifications for the replacement system were developed, a Request for Information (RFI) was issued, and based on responses from a number of potential vendors, a Request for Proposals (RFP) was distributed.

MAGGIE III Specifications

The RFP was a statement of the functions that were required on the system and the performance standards that were expected. These performance standards were based on the leading edge of development of existing hardware and software. A discussion of these standards follow.

Reliability

It is required that MAGGIE III be available as much as—if not more than—any other system on the market. It is expected that the system would be available over 99.9 percent of the time.

Capacity

The initial system must support 300 terminals operating within stated response time limits, and the architecture must allow growth to 1200 terminals without making initial hardware and software obsolete.

Expandability

It must be possible to add additional devices in increments that provide an even growth curve. In other words, it should be easy to add processor power or disk storage in increments with predictable costs; again, this must be done without making initial hardware or software obsolete.

Speed

The system is required to perform at an extremely high level of throughput. For example, the average charge-out of a book should be no more than two seconds when 300 terminals are on the system.

Mainstreamed

The hardware, operating system, programming languages, and peripheral devices should be standard off-the-shelf products. Terminals should be available from a number of vendors and should be low priced.

Vendor-Supported Housekeeping Programs

In a major departure from past practice, the decision was made to seek software for circulation, for a public access catalog, for acquisitions, and for serials from a vendor, if the program met our needs and ongoing support was available. If these programs were not available, then PPLD staff would continue to develop the programs.

In-house Enhancement

PPLD wanted to retain the ability to develop applications in-house or to add packages from other vendors such as a financial package. A report and query language was required from the computer vendor that would allow PPLD staff to interact with the applications supplied by the vendor. The system should provide a database manager and other utility programs which decrease development time significantly.

Technical Ability

It was required that the system provide the best technical features of any system available on the market and that the architecture for the system would be optimized for PPLD needs.

"User Friendliness"

Even in 1983, the staff of PPLD—and the patrons—had several years of experience with a public access catalog and had developed a number of definite prejudices. To make everyone happy, the system would have to be accurate, fast, and powerful, but simple as well. Since PPLD had over 3000 users accessing the system from their home computers, ease of use took priority. PPLD staff traveled all over the country in order to evaluate the user friendliness of the proposed systems.

CARL Expanded Ability

None of the initial proposals met the PPLD specifications, and it was only when Eyring Research Institute, Inc. proposed a system using Tandem Computer Corporation hardware, the housekeeping software from the Colorado Alliance of Research Libraries, that PPLD signed a contract, mindful of Eyring's expertise in installation, documentation, and training. In addition, Eyring agreed to incorporate some changes that were desired by PPLD staff. A detailed analysis of PPLD staff determined that the CARL system could provide some capabilities beyond the other vendors.

Database Manager

The CARL Public Access Catalog (PAC) could also serve as a database manager. The unique search routine using word, name, or browse is neutral as to content of the file. In other words, a file of clubs can be searched by word, name, or browse as easily as the catalog of books. Since the CARL system uses the MARC record, which allows variable-length fields within variable-length records, the PAC can be used to file almost anything. The system is easily explained to the average user by stating that it is as if every word and name on the catalog card is indexed, and every possible combination of the words on the catalog card can be used. Perhaps it should be labeled "the vacuum cleaner" approach to indexing since the number of access points to a specific record is in the hundreds (one possible way to calculate the number is to count the words in the record and use that number factored).

A Network System

The CARL Public Access Catalog was designed to allow the user to choose among the catalogs of all of the members of CARL, which facilitates choosing not only among different types of library's catalogs but among files as well. It is anticipated that MAGGIE will be connected directly to the CARL network, thereby providing access to over 80 percent of the titles in public and university libraries in Colorado.

A Database Manager Supervisor

Not only can the user of the CARL system switch among catalogs of different libraries, but the initial search may also be repeated in each catalog automatically. It is anticipated that a global search of all files will soon be possible.

These expanded capabilities of the Eyring system fit the needs of PPLD extremely well. Since PPLD views its mission as one of community information center and community communications center, as well as of traditional published materials center, it is essential to have a system with which to create online database systems. The ability of the system to allow in-house design and creation of new databases in a nominal amount of time places PPLD on the leading edge of agencies providing community information. A new file was designed and implemented, and data loading began in less than ten days by PPLD staff. This file, called "KWIKREF" is for miscellaneous information developed by staff research that the librarians wish to retain indexed by every word and name.

MAGGIE III Implementation

A contract was signed with Eyring Research Institute, Inc. on 29 March 1985; the hardware was delivered four days later and was installed in another two weeks. The CARL software was installed within another two weeks, and a circulation system and a public access catalog system were fully functional five months after contract signing. The hardware for MAGGIE III consists of four Tandem Non-Stop TXP 32-bit processors with four megabytes of memory each, four V8 disk drives with a capacity of 3 billion bytes, a high-speed tape drive, a high-speed printer, and the cabinetry, etc. for 300 terminals. The terminals purchased are primarily Lear Sigler ADM 12s that cost less than $600. Since the system is quite flexible on the selection of terminals, all of the old terminals from MAGGIE II are usable on the PAC. The system has exceeded performance specifications significantly. It appears that the system will handle more than 500 terminals and maintain the current response time of less than a second to charge a book. The system has been operational more than 99.9 percent of the time in the first nine months, and the public is impressed with the ease of use of the Public Access Catalog.

PPLD is extremely pleased with the implementation of the system. It has performed at levels exceeding specifications in all functions. The Community Resource Files that were present on MAGGIE II have been implemented on MAGGIE III with the significant improvement of consistent screens, terminology, and search strategy among all files. In addition, a catalog of documents created by local government agencies has been implemented. The time to develop a local database on the system appears to be only 10 percent of the time required on MAGGIE II.

A Look at the PAC

A look at the screens of the PAC will illustrate the excellent user friendliness of the CARL PAC. It should be noted that the screens have been merged into exhibits in order to make the presentation more compact. The text is as it appears; I have simply eliminated the majority of the blank lines on the CRT screen.

Exhibit 1 shows the initial screen seen on PAC terminals in library facilities and on home computers linked via dialup. The HELP choice at this point simply explains the contents of each database.

Exhibit 2 shows the screen that appears after selecting number 1, the ON-LINE CATALOG. This screen explains the type of searches that are available. Next on the screen is what appears after selecting *W* for Word search. Different examples are provided for each file in order to be more relevant to the user. The next several lines show the result of the user

```
(The  computer  screens  have been merged in order  to  provide  a
savings in space.   The text you see was downloaded directly into
the wordprocessor from the PAC)


WELCOME TO THE COMPUTER CATALOG OF LIBRARY HOLDINGS
(version 50)

                    PIKES PEAK LIBRARY DISTRICT

          A project to the Eyring Research Institute and the
            Colorado Alliance of Research Libraries (CARL)

PRESS (RETURN) TO START PROGRAM:

PAC WORKING...
Your first step is to select the LIBRARY whose catalog you wish to consult.

Catalogs are currently available for:

1.    ON-LINE CATALOG
2.    CALENDAR
3.    AGENCY
4.    CLUB
5.    COURSES
6.    LOCAL DOCUMENTS
7.    KWIKREF
8.    HELP...
9.    DAYCARE


TYPE the NUMBER of the library you wish to search, and
press the (RETURN) key.

ENTER NUMBER:1
WORKING...
09/22/86
```

Exhibit 1

entering the words *transportation planning*. The sequence of the words is irrelevant and the user may enter word stems if he or she is not sure of the complete word. For example, *plan* might have been used to expand the search to include *plan* and *plans*. The catalog contains 175 titles under the term *transportation* and ten titles under *transportation* + *planning*. At this stage the user asks for a list of those ten items.

Exhibit 3 shows the list of the first seven hits on *transportation planning*. When the user selects one of the numbers the screen containing the record is displayed. This screen is the complete MARC record, with call number, facility location, and status. At the bottom of the screen the system provides the option of repeating the search on another database. By entering *S*, the user may return to the screen to choose another database. When the user enters the number 7, the search is repeated on the LOCAL DOCUMENTS database as shown on exhibits 4, 5, and 6. Because of the large number of documents in the database on *transportation planning*, it is

```
02:40 P.M.         SELECTED CATALOG :     ON-LINE CATALOG

            The computer can find books by NAME or by WORD

            NAMES can be authors, editors, or names of
            persons or institutions written about the book

            WORDS can be words from the title, or subjects,
            concepts, ideas, dates etc.

            You may also BROWSE by TITLE, CALL NUMBER, OR SERIES.

            Enter    N    for   NAME search
                     W    for   WORD search
                     B    to    BROWSE by title, call number or series
                     S    to    STOP or SWITCH to another Library catalog


     There is also a quick search  -- type QS for details

            Type the letter for the kind of search you want,
            and end each line you type by pressing <RETURN>
                    SELECTED CATALOG :     ON-LINE CATALOG


ENTER  COMMAND>>  W
 WORKING...



REMEMBER -- WORDS can be words from the title, or can be subjects,
concepts, ideas, dates, etc.

           for example --  GONE WITH THE WIND
                           SILVER MINING COLORADO
                           BEHAVIOR  MODIFICATION
enter word or words (no more than one line, please)
separated by spaces and press <RETURN>.

>TRANSPORTATION PLANNING
 WORKING...
TRANSPORTATION 00174TITLES
TRANSPORTATION + PLANNING 00010TITLES

For the   00010 items that have
TRANSPORTATION + PLANNING
Press <RETURN>, or type <Q>UIT for a new search.
 WORKING...
```

Exhibit 2

necessary to add additional terms to narrow the search. The user can scan short entries for all items if he/she wishes.

Exhibits 7 and 8 show the search *transportation planning* repeated on the Agency Database with the list of short entries in that file. The system provides an interesting information resource scanning facility. I doubt that most transportation planners would think of the Girls Club or day-care centers as entities involved in transportation, but the file shows that is certainly the case.

```
PREPARING YOUR DISPLAY -- HOLD ON...
  1 Meyer michael d                                  PPLD see record  1984
  Urban transportation planning :  a decision-orie 388.4068 M613u

  2                                                  PPLD see record  1983
  Metropolitan transportation planning               388.4068  M594

  3 Foster mark s                                    PPLD see record  1981
  From streetcar to superhighway :  american city  771.70973 F756f

  4 PIKES PEAK                                       PPLD see record  1979
  Notebook for first transportation planning works 711.7     P6369N

  5                                                  PPLD see record  1976
  Out of cars, into transit :  the urban transport 388.40973 094

  6 Citizens' goals colo                             PPLD see record  1976
  [citizens' goals background papers]                307.760978C581B

  7 GRAY,GEORG                                       PPLD see record  - nd
  Public transortation: planning, operations, and  388.4     G779P


 <RETURN> TO CONTINUE DISPLAY

ENTER <LINE NUMBER> TO DISPLAY FULL RECORD

<Q>UIT FOR NEW SEARCH 3
WORKING...
-----------------------------------------------------------------
AUTHOR(s):        Foster, Mark S.
TITLE(s):         From streetcar to superhighway :  American city planners
                    and urban transportation, 1900-1940 / Mark S. Foster.
                  Philadelphia :  Temple University Press,  c1981.
                  xiv, 246 p. :  ill. ;  24 cm.
                  Technology and urban growth
                  Includes index.
                  Bibliography: p. 235-237.

OTHER ENTRIES:    Urban transportation policy  United States  History.
                  City planning  United States  History.

CALL #: 771.70973 F756f                      LOCN:  PENROS
STATUS:  Not checked out --

<RETURN> to continue, <Q>UIT for a new search, or <R> to REPEAT this display
Q

You began with a W search on:
TRANSPORTATION PLANNING
Type S to try your search in another catalog, or
     R to repeat your search in ON-LINE CATALOG or
     <RETURN> for a new search
```

Exhibit 3

```
Your initial search was:

   TRANSPORTATION PLANNING

Select the catalog you wish to try next:
1.    ON-LINE CATALOG
2.    CALENDAR
3.    AGENCY
4.    CLUB
5.    COURSES
6.    LOCAL DOCUMENTS
7.    KWIKREF
8.    HELP...
9.    DAYCARE

TYPE the NUMBER of the library you wish to search, and
press the <RETURN> key.

ENTER NUMBER:6
WORKING...
this takes a sec...
SELECTED CATALOG:    LOCAL DOCUMENTSB
TRANSPORTATION 01134TITLES
TRANSPORTATION + PLANNING 00761TITLES

For the   (00761) items that have
TRANSPORTATION + PLANNING
Press <RETURN>, or type <Q>UIT for a new search.
 WORKING...


TRANSPORTATION + PLANNING 00761TITLES


You may make your search more specific  (and reduce
the size of the list)  by adding another word
to your search. The result will be items in
your current list that also contain the new
word.

 to ADD a new word, enter it,
 <D>ISPLAY to see the current list, or
 <Q>UIT for a new search:

NEW WORD(S): COLORADO
 WORKING...
TRANSPORTATION + PLANNING  + COLORADO 00137TITLES

You now have: TRANSPORTATION + PLANNING  + COLORADO 00137TITLES
```

Exhibit 4

```
You may make your search more specific (and reduce
the size of the list) by adding another word
to your search.  The result will be items in
your current list that also contain the new
word.
 to  ADD a new word, enter it,
 <D>ISPLAY to see the current list, or
 <Q>UIT for a new search:

NEW WORD(S): SPRINGS
 WORKING...
TRANSPORTATION + PLANNING  + COLORADO + SPRINGS 00081TITLES


You now have: TRANSPORTATION + PLANNING  + COLORADO + SPRINGS 00081TITLES

You may make your search more specific  (and reduce
the size of the list)  by adding another word
to your search. The result will be items in
your current list that also contain the new
word.

 to ADD a new word, enter it,
 <D>ISPLAY to see the current list, or
 <Q>UIT for a new search:


NEW WORD(S): HIGHWAYS
 WORKING...
TRANSPORTATION + PLANNING  + COLORADO + SPRINGS + HIGHWAYS 00009TITLES

You now have:  TRANSPORATION + PLANNING + COLORADO + SPRINGS + HIGHWAYS
00009TITLES

You may make your search more specific (and reduce
the size of the list)  by adding another word
to your search. The result will be items in
your current list that also contain the new
word.
 to ADD a new word, enter it,
 <D>ISPLAY to see the current list, or
 <Q>UIT for a new search;

NEW WORD(S): D
```

Exhibit 5

The Future of MAGGIE'S PLACE

There are several new databases on the drawing boards and, based on the implementation of the current databases, it appears that new ones will be created on a regular and frequent schedule. Plans may change depending on circumstances, but at present the several databases are planned.

Facts

Pierian Press has provided a demonstration tape of the data contained in its serial entitled *A Matter of Fact.* It appears that loading this data into the PAC will be relatively routine. Pierian Press will send out updated

```
PREPARING YOUR DISPLAY -- HOLD ON...
  1 Agency colorado depa                            PPLD PENROS LOHIST - nd
  Title: u.s. highway 24 bypass colorado springs,  CRDO+HWY/EI-T83(1976)

  2 Agency pikes peak ar                            PPLD PENROS LOHIST - nd
  Title: study of access routes to peterson field  REG+PPACG/SP-P37(1967)

  3 Agency planning divi                            PPLD PENROS LOHIST - nd
  Title: transportation plan, city of colorado spr CS:CD-PL/SP-T61(1986)

  4 Agency ridefinders t                            PPLD PENROS LOHIST - nd
  Title: a proposal for alternative transportation ORG+RIDE/SP-G16(1984)

  5 Agency department of                            PPLD PENROS LOHIST - nd
  Title: traffic and preliminary engineering study CRDO:HWY/SP-G16(1986)

  6 Agency pikes peak ar                            PPLD PENROS LOHIST - nd
  Title: colorado springs long-range plan update s REG:PPACG/SP-T851(1984)U

  7 Agency colorado depa                            PPLD PENROS LOHIST - nd
  Title: widefield, el paso county; draft environm CRDO+HWY/EI-W42(1971)

 <RETURN> TO CONTINUE DISPLAY

ENTER <LINE NUMBER> TO DISPLAY FULL RECORD
<Q>UIT FOR NEW SEARCH 7

WORKING...
----------------------------------------------------------------
AUTHOR(s):       AGENCY: COLORADO DEPARTMENT OF HIGHWAYS
TITLE(s):        TITLE: WIDEFIELD, EL PASO COUNTY; DRAFT ENVIRONMENTAL
                   IMPACT STATEMENT ADMINISTRATIVE ACTION

                 PUB DATE: SEPTEMBER 30, 1971
                 ABSTRACT: THIS STATEMENT DISCUSSES THE PROPOSED ROUTE FOR
                   THE EXTENSION OF STATE HIGHWAY 16. THE EXTENSION DISCUSSED
                   CONTINUES HIGHWAY 16 EASTERLY TO INTERSECT THE EXTENSION OF
                   MARKSHEFFEL ROAD. BOTH ROUTES HAVE BEEN ACCEPTED AS PART OF
                   THE COLORADO SPRINGS METROPOLITAN AREA TRANSPORTATION STUDY
                   FOR THE COLORADO SPRINGS AREA.

OTHER ENTRIES:   KEY WORDS: TRAFFIC - PLANNING; HIGHWAYS; SOCIO-ECONOMIC
                   ANALYSES; MASS TRANSPORTATION; CRDO+HWY/EI-W42(1971)
                 DOC TYPE: ENVIRONMENTAL IMPACT STATEMENT
                 GEOG AREA: WIDEFIELD
                 FEATURES: MAPS, PHOTOGRAPHS, DIAGRAMS
                 GOVT LEVEL: STATE
                 XDONOR: XPPACGLIBRARY

CALL #:  CRDO+HWY/EI-T83(1976)      LIBRARY:      PENROS LOHIST
```

Exhibit 6

tapes as a subscription with fixed prices. By the time this article appears in print, the users of PPLD will be routinely searching the data contained in the "FACTS" database by any word or name. It should be very popular with home users since there will be tens of thousands of facts with citations of sources available in their own homes.

Reviews

The editor of Pierian Press has indicated that they would like to create a database consisting of most of the book reviews contained in magazines

```
<RETURN> to continue, <Q>UIT for a new search, or <R> to REPEAT
this display
Q
You began with a W search on:


TRANSPORTATION PLANNING

Type S to try your search in another catalog, or
     R to repeat your search in    LOCAL DOCUMENTS or
     <RETURN> for a new search:S
Your initial search was:

   TRANSPORTATION PLANNING

Select the catalog you wish to try next:
1.    ON-LINE CATALOG
2.    CALENDAR
3.    AGENCY
4.    CLUB
5.    COURSES
6.    LOCAL DOCUMENTS
7.    KWIKREF
8.    HELP...
9.    DAYCARE

TYPE the NUMBER of the library you wish to search, and
press the <RETURN> key.

ENTER NUMBER:3
WORKING...
this takes a sec...
SELECTED CATALOG:    AGENCY
TRANSPORTATION 00027TITLES
TRANSPORTATION + PLANNING 00002TITLES


PREPARING YOUR DISPLAY -- HOLD ON...

                                                    PPLD
1  Pikes peak area council of governments           AGENCY FILE


                                                    PPLD
2  Downtown colorado springs, inc.                  AGENCY FILE


ALL ITEMS HAVE BEEN DISPLAYED.
ENTER <LINE NUMBER> TO DISPLAY FULL RECORD
<Q>UIT FOR NEW SEARCH 1
```

Exhibit 7

```
WORKING...

TITLE(s):         PIKES PEAK AREA COUNCIL OF GOVERNMENTS
                  ADDRESS:  27 E. VERMIJO, 5TH FLOOR
                  CITY,ST,ZIP:  COLORADO SPRINGS, CO 80903
                  HOURS:  8-5 M-F
                  TELEPHONE:  471-7080
                  PARENT ORG:  PPACG
                  DIRECTOR:  DAVID L. PETERSON
                  CONTACT:  DIAN SUKALSKI

OTHER ENTRIES: FUNCTION:  REGIONAL PLANNING AGENCY FOR EL PASO, PARK AND
                  TELLER COUNTIES; REGIONAL CLEARINGHOUSE FOR FEDERAL
                  FINANCIAL ASSISTANCE (A-95 REVIEW); OFFICIAL CENSUS DATA
                  USER CENTER PROVIDING CENSUS DATA, MAPS, STATISTICS, ON
                  HOUSING AND POPULATION; PROVIDE PLANNING SERVICES IN
                  TRANSPORTATION, AGING, ENVIRONMENTAL QUALITY, HOUSING,
                  STATISTICS/ECONOMIC FORECASTS, ETC; MAINTAIN EXTENSIVE MAP
                  LIBRARY AND PLANNING LIBRARY.  HANDICAPPED ACCESS.
               ELIGIBILITY:  LOCAL GOVERNMENTS, PRIVATE ACCESS.
               APPLICATION:  PHONE, WALK IN
               KEYWORDS:    PLANNING        GOVERNMENT      LOCAL
               DATE ENTERED:  06-Mar-81
               DATE UPDATED:  08-Apr-86
               AGENCY:  PIKES PEAK AREA COUNCIL OF GOVERNMENTS

CALL #: AGENCY FILE               LIBRARY:

----------------------------------------------------------------
<RETURN> to continue, <Q>UIT for a new search, or <R> to REPEAT this
display
You began with a W search on:
```

Exhibit 8

throughout the United States. They would supply monthly update tapes of PPLD so that current reviews would be available in the PAC. It would be an extremely valuable source for librarians involved in book selection, and the public would be able to view a number of reviews of a book prior to selecting its reading material.

Voter Information

The Pikes Peak area has many representative government divisions in which the citizens need to participate through elections. There is a plethora of districts as well as the cities and the county. El Paso County alone has twenty-six school districts. The polling places for elections can be difficult to find.

City Code

The city attorney for Colorado Springs has asked PPLD to assess the feasibility of providing dial-up access to the city legal code. This code is modified practically every time the city council meets and it is very expensive to reprint the code every two weeks.

LINK

There are many people willing to teach those who want individualized instruction on almost any topic. The LINK file would simply be a directory of people and their talents.

Conclusion

User friendliness must move beyond traditional, esoteric cataloging practice as embodied in the card catalog. It means simple commands that are very powerful, operating on multiple databases that are relevant to the local user community. There is little question at PPLD that the methods of Plato and Aristotle exist side by side in the online catalog. MAGGIE'S PLACE has entered a new era.

CHRISTINE L. BORGMAN
Graduate School of Library and Information Science
University of California—Los Angeles

Toward a Definition of User Friendly: A Psychological Perspective

Introduction

"User friendly" is one of those valuable concepts that has become such an overworked phrase that it has lost much of its meaning. As Meads notes, "[t]he forced grin of user friendliness becomes a mask for lack of capability, insufficient performance, costly maintenance, or a collection of mis-fitting components."[1] User friendly is not merely the addition of high tech hardware such as a mouse, icons, or three-dimensional graphics.

What does *user friendly* mean? First, consider a dictionary definition. Webster's defines *user* simply as "one who uses."[2] *Friendly* is defined as "of, relating to, or befitting a friend: as *a*: showing kindly interest and goodwill *b*: not hostile *c*: inclined to favor *d*: comforting, cheerful."[3] We infer that "user friendly" suggests an entity that is warm and comforting to the one who uses it.

Matthews and Williams defined a "user friendly index" for information systems as a nine-point scale, ranging from "user intimate" at the top to "user vicious" at the bottom, with "user oriented" as a midpoint.[4] They have followed the same line as Webster's, considering "user friendly" as being kind—or at least the inverse of hostile—to the one who uses.

Meads takes the definition of user friendly further by stating three requirements. The first is that the system is *cooperative*—it provides active assistance during the task and makes its actions clear and obvious. Second, the user friendly system is *preventive*—it acknowledges that people make mistakes by preventing those mistakes to the extent possible and by providing backout and recovery procedures. Third, the friendly system is *conducive*—it is reliable, predictable, and assists rather than controls the user.[5]

Meads's three requirements can be combined into the one attribute transparency, a commonly used term from computer science. If a system is transparent to the user, it means that the user is looking through the system to the task being accomplished and not focusing on the system itself. A transparent system is one that supports and simplifies a task rather than becoming a task in and of itself. This paper will discuss current research on information systems that has the implicit goal of making systems more user friendly and that is being conducted from a psychological perspective.

The Human as a Unit of Analysis

The interaction of humans with computers can be studied at multiple levels of analysis. Here we are concerned with the psychology of the user, which is roughly a mid level unit of analysis. By psychology we mean the study of human behavior—i.e., mental and behavioral characteristics as they apply to the use of computers. The research done in the area is largely based on the theories of cognitive psychology. Studies are of the individual user as representing the larger body of users.

Human-computer interaction can be studied at both lower and higher levels of analysis than that of the individual user. At a lower level would be the human factors studies that focus on anthropomorphic dimensions of the human: fitting the keyboard size and layout to the average human hand, designing workstations with the proper dimensions for human comfort, screen displays that minimize glare and eyestrain, and so on.

At a higher level of analysis than the individual is the study of the organization or the social group response to the use of computers. The way in which people use computers is affected by the way in which the systems are introduced, their motivation to use them, the training provided, the threats to the current job, and changes in task and work structure.

All of these levels must be studied to provide a full picture of human use of computers and hence of friendliness. However, they cannot all be studied at once. In this paper we confine ourselves to the study of the individual user.

A Psychological Perspective

Researchers in academic departments of psychology, communication, computer science, and library and information science, as well as industrial researchers, have been applying both psychological theory and method to the study of human interaction with computers. In addition, psychologists have used the study of human behavior with interactive systems as a test-bed for developing theory and method.

The remainder of the paper will cover two distinct bodies of research. First we cover psychological theories that have been applied directly to interactive computer systems. Some theories already have been applied to information systems; others are better proven elsewhere but have potential for use in this domain.

The second body of research to be addressed is studies done to characterize behavior on information retrieval systems—both online catalogs and bibliographic retrieval systems—that is not driven by theory. Rather, it is pretheoretical, gathering data that may lead to theory development later. This body of research utilizes research methods drawn from psychology and other social sciences. We will focus specifically on studies of error behavior because errors interfere with usage and hence with transparency.

APPLICATIONS OF PSYCHOLOGICAL THEORY

Three theories will be considered here, each of which is general and has been applied to other information technologies. The first is that of mental models, an attempt to describe the learning and problem-solving processes involved in the use of computer systems. Second is that of information processing models, an attempt to build discrete quantitative models of interactive behavior. The third theory considered is individual differences, an attempt to explain variance in performance and interaction style by personality and demographic characteristics.

Mental Models

The mental models theory, drawn from cognitive psychology, is perhaps the most appealing theory for the study of human behavior on information systems.[6] Although it has not yet been applied widely to retrieval systems, the research to date holds considerable promise for both design and training.

Psychological Research on Mental Models

Research in learning theory in various contexts has shown that people tend to build hypotheses as part of problem solving. When a person approaches a new task, whether it's fixing a toaster or a carburetor, solving a math problem, or learning a text editor, he or she tends to gather information from the context of the task. The information might be drawn from a manual, from watching other people, from prior knowledge, or from the response of the problem to the user's actions. As the user/problem solver takes an action—such as turning a screw, writing an equation, or entering a command—the problem changes and the result is observed. From all of these sources the user makes further hypotheses about how the entity or problem works and about why it is responding in a particular way. Evidence from actions is taken as supporting or negating the hypotheses made and the hypotheses are refined accordingly with the user taking more actions until the task is completed or abandoned.

All these hypotheses and actions fit together into a "mental model" of how the entity works. The mental model starts out fuzzy and becomes more clearly defined with experience. It is important to note that the user/problem solver is not necessarily aware that he or she has or is applying such a model. The model is part of the problem-solving process and usually is not a conscious effort.

The ability to develop a mental model is a valuable intellectual skill and one that is very helpful when the information applied to the problem is correct and when the hypotheses are correctly interpreted or revised. Unfortunately this is not always the case. A person may not gather enough information about the task first (read the instructions, assess the nature of the problem), or he or she may start with incorrect assumptions such as that it works like some other entity previously seen or that the problem is something other than it actually is. For example, people often assume that a text editor works much more like a typewriter than it actually does or that an online catalog is more like a card catalog than is actually the case. To complicate life further, people often interpret the results of their actions as supporting their hypotheses whether or not they do indeed.[7]

The theory suggests that people can be trained with a conceptual model of the system from which they can draw a mental model that is compatible with their own thinking processes. The research design typically applied in mental models studies is to assign subjects to two groups, one trained with a conceptual model of the system and one trained with a procedural set of instructions (no framework; just "first do this, then do this...."). The underlying hypothesis is that those trained with a conceptual model will develop a mental model and will perform better on the tasks, and those trained only with procedures either will develop an incorrect model or will not develop a model at all.

A further hypothesis is that having a mental model is not as important for the simple tasks that can be accomplished with one or two predefined procedures as it is for more complex tasks that involve multiple procedures or extrapolation from basic procedures.[8]

Applications to Information Systems

The first study to test the mental models theory on retrieval systems compared the two training methods on a Boolean-logic-based online catalog of OCLC records.[9] As predicted, it was found that on simple tasks there was no difference in performance (number of items correct) based on training, but on complex tasks, those trained with a conceptual model of the system got more items correct and exhibited different patterns of interaction with the system than those trained procedurally.

The only other study of mental models and information systems identified to date is a master's thesis from the University of Chicago done by Jean Dickson.[10] Her study was not experimental; rather, she attempted to infer a mental model from the monitoring record of user behavior on NOTIS, the online catalog at Northwestern University. Dickson looked specifically at the errors in author and title searches that resulted in no hits and concluded that users applied different mental models from those applied to a card catalog because they searched differently. Her most striking examples were the frequency of errors due to entering authors with given name first (12.6 percent of no-hit author searches) and due to the inclusion of initial articles in title searches (10.1 percent of no-match title searches), neither of which would be appropriate behavior in a card catalog. Other explanations exist for these behaviors, but the data do suggest that users make incorrect hypotheses about the system.

Information Processing Models

Psychological Research on Information Processing Models

An information processing model is an attempt to break down human tasks into discrete physical and cognitive actions and to assign probabilities of occurrence and performance times to these actions. The model allows task behavior to be calculated and predicted. The computed performance times and patterns can be used to compare methods of performing a given task. The best known of these models are the GOMS (Goals, Operators, Methods, and Selection rules) and keystroke models of Card, Moran, and Newell.[11] The GOMS model predicts human behavior on a specific task in terms of the user's goals, operators, methods, and selection rules. The model was developed using manuscript editing tasks. In this context, Card and his colleagues have achieved roughly 90 percent accuracy in predicting behavior sequences and 33 percent accuracy in predicting time required for modifications.

The keystroke model is more discrete and predicts time to perform a given task as a linear sum of four physical and one mental operators. In text-editing tests, Card's research team modeled behavior with a 21 percent error rate.

These models are useful for comparing features for implementation in designing a system. They have been used for comparisons such as determining whether a control character sequence is better for an editing function than a function key or whether a mouse is better than a joystick for pointing to objects on a screen.[12]

The information processing models on a task are built by training people until they are expert which may take thousands of repetitions. It has been reasonably successful in developing text editing systems which are well-suited to expert, highly repetitive behavior. The models also are being used to advance the information processing theories of cognitive psychology.

Applications to Information Systems

The information processing models have not yet been applied to the design of information retrieval systems. They may be helpful for determining the best use of command sequences in terms of making frequently used actions most accessible, minimizing confusion among actions, and so on.

Overall, the information processing models are less applicable to information retrieval systems than to text editing because the task does not lend itself as well to expert behavior. The information retrieval task is much less clearly defined, requiring heuristic thinking and continual reevaluation of the task. Further, few users of information retrieval systems use them in a production, expert mode. The vast majority use the systems too infrequently to achieve the expert behavior on which the information processing models are based.

Individual Differences

Psychological Research in Individual Differences

Most of the psychological research on human interaction with interactive systems comes from the area of cognitive psychology which is based on the "information processing model" paradigm alluded to earlier. The theory which underlies much of current cognitive research attempts to reduce human behavior to information inputs, processes, and outputs. The intent is to identify fundamental characteristics across all people that can be used to predict behavior. The information processing theorists do not acknowledge differences among people. Rather, they treat such differences as "random variance."

Another branch of psychology is specifically interested in that "random variance." Those in the area of "correlational" or "differential" psychology look for variance in behavior that occurs naturally and then seek factors that differentiate among individuals or groups. Their intent is to identify causal, or at least associative, relationships after the fact.

The differential psychology researchers have determined that some people have an easier time using information technologies than others—including information retrieval systems, text editors, and programming languages. Once the fact has been established that a range of behavior exists, the method is to analyze the behavior of a group of people on the task, capturing data on as many related factors as are hypothesized to be responsible for the differences.

In text editing studies, researchers have found that age and spatial memory are important factors.[13] Those who are younger and who have the best spatial memory capabilities perform best on text editors.

Similarly, researchers have found consistent variance in those who are professional programmers, finding that they fall into a consistent style of processing—more thinking than feeling, more intuitive than sensing.[14] Those who perform best in introductory programming courses also take more science and math courses, score better on general achievement tests (math and verbal), and get higher grades.[15]

Applications to Information Systems

Studies of user behavior on both bibliographic retrieval systems and online catalogs long have found wide variance in usage patterns even when the same system and database are used.[16] In summarizing the characteristics of the "average" search across multiple studies, Fenichel reports broad ranges in reported means for variables such as number of descriptors searched, commands used, connect time, retrieved references, recall, precision, and unit cost.[17] Only recently have researchers begun to identify systematically the sources of some of the variance observed.

Amount of experience with the system is the variable most commonly studied in identifying performance differences. Fenichel was able to determine only that novices (low database experience and low searching experience) searched more slowly and made more errors than experienced searchers.[18]

Penniman, in monitoring studies, found that frequent searchers of the NLM Medline system used about the same number of single terms and displays in a search as did infrequent searchers but twice as many advanced term search entries and half again as many Boolean searches.[19] Moderately frequent searchers used more of all types of commands than infrequent users.

Three dissertations have explored the personality differences that may underlie searching performance on bibliographic retrieval systems. Brindle studied the relationship between cognitive style and search performance in a field experiment but found few significant differences.[20] Bellardo studied graduate library school students who had just completed a course in online searching, testing them on two measures of creativity and one measure of personality and obtained their Graduate Record Exam (GRE) scores. Bellardo attempted to correlate these measures with search performance (precision and recall) but was unable to explain much of the variance. However, she did find a significant ($p < .05$) correlation between search performance and GRE quantitative scores but no correlation with GRE verbal scores.[21]

In a field experiment, Woelfl tested skilled NLM Medline searchers on inductive and deductive reasoning and learning style. Woelfl found that searchers clustered strongly in one learning style (high active, high abstract). Overall, the cognitive attributes affected the search process but not search results.[22]

As with other types of information retrieval systems, we find a wide range in skills among online catalog users. Monitoring studies have identified high variance in the types of searches performed, in the length of searches, and in the patterns of errors.[23] Each of these were unobtrusive field studies and did not collect any data on individual users that could be compared to the search pattern data. Survey data of the same population found a comparable range of user-reported success and satisfaction levels in system use and a broad range of user background characteristics.[24]

Borgman found significant differences in the ability to pass a benchmark test of information retrieval skills by academic major. Those who failed the test were predominantly social science and humanities majors while those passing the test were science and engineering majors ($p < 0.0001$). Prior computer experience was controlled (subjects had no information retrieval experience and at most two programming courses).[25]

Based on the earlier discussed results, Borgman is pursuing the hypothesis that academic major is a gross measure of individual differences and is probably a surrogate for other characteristics that are associated with major.[26] Preliminary results of a study incorporating personality tests used by Woelfl and demographic characteristics identified in studies of programming aptitude indicate that engineering majors cluster strongly around personality characteristics associated with both information retrieval and programming, while English and psychology majors show either no pattern or one opposite that of engineering majors.[27]

ERROR BEHAVIOR

The study of error behavior is crucial to the issues of system transparency. If a system is transparent, it will support and simplify a task—not become a task in itself—and be congruent with the user's thinking style and workflow. The difficulty is in measuring these indicators of transparency. We find usually that it is easier to gather evidence on when a system is *not* working well than on when it *is*. Thus, we study user errors and problems.

User errors and problems with information retrieval systems can be divided into two categories: those encountered with the mechanical aspects of searching (typos, incorrect commands, etc.) and those with the conceptual aspects (controlling the interaction, achieving useful results, etc.).

By identifying errors in the mechanical aspects, we can identify poorly engineered system factors that may be increasing the likelihood of certain types of errors. Identifying the most common errors can lead to isolating nonintuitive command sequences, misleading displays, and other unfriendly aspects of a system.

Similarly, by identifying poor levels of searching performance (low recall and precision, inefficient use of commands, etc.), we can determine ways in which the system interferes with the natural flow of problem solving (retrieving information) and the points at which it fails to be congruent with thinking style and workflow. It also allows us to identify misconceptions about the systems thereby understanding better how people are interpreting system actions and internalizing them into their behavior. With such knowledge both the design of systems and training for them can be improved.

The causes of the errors and problems identified by studying user behavior can only be inferred, of course. But the evidence will result in hypotheses about the sources of the behavior that can be taken to the laboratory for further study.

The discussion here is intended to provide only an introduction to the kinds of studies that can be done to identify user problems with systems. For a fuller discussion of these results and their implications, the reader is referred to Borgman[28] (the applications of psychological theory are discussed at length in another paper by Borgman[29]).

Problems with Mechanical Aspects of Searching

Bibliographic Retrieval Systems

Problems with the mechanical aspects of searching have not proven to be a major barrier to the use of bibliographic retrieval systems, although several studies have found that they are a barrier for very inexperienced and infrequent users.[30] Fenichel, in an experiment capturing printed search protocols, found that both moderately experienced and very experienced searchers made significantly fewer nontypographical errors per search than did novices although the overall number of errors was small (2.8 per search for novices).[31]

Defining errors only as erasures, Penniman found an average of 8 percent of user actions as errors.[32] Tolle and Hah, using the same definition in a monitoring study of the NLM CATLINE database, also found an average error rate of 8 percent.[33]

Online Catalogs

Mechanical problems have been particularly evident in monitoring studies of online catalogs. Tolle found that errors were not isolated.[34] Instead they tended to occur in clusters; once an error was made the next transaction was likely to be an error as well. In the SCORPIO system of the Library of Congress, given that an error was made, the likelihood that the next command was an error was 59.8 percent; for the SULIRS system at Syracuse University, it was 28.6 percent; for the LCS system at the Ohio State University it was 33.3 percent. Errors were defined in SCORPIO as unrecognizable search commands; in SULIRS as an unrecognizable command, an incorrectly formatted command, or an invalid item number; in LCS as partially or fully unrecognizable commands. Data from these studies also indicate that users tend to quit immediately after receiving an error message.

In a monitoring study of the Ohio State University (LCS) online catalog, Borgman defined two types of errors: logical errors or commands that could be partially recognized by the system and typing errors or commands that could not be recognized at all. Errors were roughly equally divided between the two types. Total errors averaged 13.3 percent of all user commands; 12.2 percent of all user sessions studied consisted entirely of errors.[35]

Dickson[36] and Taylor[37] analyzed the monitoring record of search input on the NOTIS system that resulted in no matches on known-item searches. Dickson found that 37 percent of all title searches and 23 percent of all author searches resulted in no matches. She determined that 39.5 percent of the no-match title searches and 51.3 percent of the no-match author searches were for records that existed in the database and were not found due to user errors in searching. Of the errors in title searches, 15 percent could be attributed to typos or misspellings; the remaining errors were conceptual in nature.

Taylor found that only 22.4 percent of the no-match author searches could be determined to be good author names that were not in the database; the remaining 77.6 percent could have been for records actually in the database. She was able to attribute 22.1 percent of the no-match author searches to misspelled words.[38]

Conceptual Aspects of Searching

Bibliographic Retrieval Systems

While problems with system mechanics are rare for both experienced and inexperienced searchers of bibliographic retrieval systems, many studies have identified significant problems with search strategy and output performance.[39] Experiments using transcripts of search behavior have shown that searchers often miss obvious synonyms or fail to pursue strategies likely to be productive.[40] Similarly, searchers often fail to take advantage of the interactive capabilities of the system.

In a survey comparing searching problems to prior training, Wanger et al. found that most respondents said they had difficulty in developing search strategies "some" (47 percent) or "most" (8 percent) of the time and 36 percent said they had difficulty in making relevance judgments "some" of the time.[41]

Perhaps as a consequence of relying primarily on simple search techniques, recall scores are often relatively low even when comprehensive bibliographies were requested.[42] In reviewing studies that computed recall measures (using a variety of research methods), Fenichel shows that average recall ranges from a low of 24 percent (novices only; 41 percent average minimum recall in other cases) to a high of 61 percent. Average precision in the same set of studies ranged from 17 percent to 81 percent.[43]

Online Catalogs

The online catalog studies also have identified many problems with the conceptual aspects of searching, although they have focused more on problems related to misunderstanding of system features than to achieving high levels of performance. Similar to Fenichel's findings,[44] survey data indicate that online catalog users rarely ventured beyond a minimal set of system features. The majority of searches were simple, specifying only one field or data type to be searched; the advanced search features were rarely used; even when systems included the feature of scanning lists of index terms or headings, users didn't utilize the feature unless "forced" to do so.[45]

Survey respondents also indicated that they had problems with several of the conceptual aspects of searching, including increasing search results when too little (or nothing) is retrieved, reducing search results when too much is retrieved, and use of truncation. Users reported that they experienced a lack of control over the search process and that they found many of the codes and abbreviations in the displays confusing.[46]

In assessing problems with specific types of searching, the survey found that subject searching was the most problematic area. Users indicated that they had problems both with performing the subject search and with identifying the right subject terms. In several monitoring studies reviewed by Markey,[47] no-match subject searches range from a low of 35 percent on MELVYL[48] to a high of 57 percent in the BACS system.[49]

In the monitoring study conducted by Dickson, no-match searches could be attributed to misunderstanding the search structure, such as inclusion of initial articles (10.1 percent of no-match title searches), wrong name order (12.6 percent of no-match author searches), and the wrong forename or the incorrect inclusion of a middle initial (9.9 percent of the no-match author searches).[50] Taylor found that 16.7 percent of no-match author searches were due to putting the forename first, another 5.6 percent to the incorrect use of a middle initial, and 5.7 percent were due to searching title or subject terms in the author field.[51]

CONCLUSIONS

We have discussed the applications of psychological theory to the design of information systems—including mental models, information processing models, and individual differences—and studies of error behavior on both bibliographic retrieval systems and online catalogs. What does all of this imply for making systems more user friendly or transparent?

Implications of Psychological Theory

The results of the mental models research suggest that systems are easiest to use when they are designed around a consistent conceptual model that is readily recognizable by the user. Further, the training and instructions for the system should reinforce the model. Status indicators on the display should indicate the current location in the system, the immediately previous location, and options for the next location. All of these data are helpful in providing a comfortable framework for system use. A transparent system, in terms of a mental model, is one whose conceptual framework is readily adopted by the user, making the system simply a tool to support the task and not a task in itself.

The information processing models have less direct implications for user friendly systems design. They suggest that user actions can be quantified into a string of additive variables, including reaction time, keystroke time, and mental processing time. Therefore, through system evaluation and basic research, we should continue to seek some underlying fundamental characteristics of information retrieval behavior. The practical results of information processing models' research probably are further away from implementation than are the results of other research paths.

The individual differences research suggests is that different people approach systems in different ways, learn at different rates, and prefer different types of training and interfaces. The first step in implementing the results of individual differences research is to acknowledge that the differences exist. When user populations are small or otherwise well-defined, it may be possible to identify common characteristics (e.g., computing knowledge, retrieval knowledge, subject expertise) and tailor systems accordingly.[52] When user populations are diverse and ill-defined (as is the case with most populations of public and academic library clientele), individual differences can be acknowledged by providing multiple forms of interfaces (e.g., menu and command) and by offering multiple forms of training (e.g., classroom training, computer-assisted instruction, printed materials). The provision of options such as these, while not allowing precise tailoring to each individual, does allow users to make choices among the interface styles and training methods with which they are most comfortable.

Error Behavior and Transparency

A review of the research on error behavior suggests that users have problems with both the mechanical and the conceptual aspects of searching information retrieval systems and that the problems occur on both bibliographic retrieval systems and online catalogs. We are beginning to identify some of the problematic factors, although they vary by system. We do know that subject searching tends to be the most problematic type of search in most systems, however, and a candidate for closer study. Another common factor is the tendency to utilize only a subset of commands, not taking advantage of the more sophisticated searching features. We need to determine if the higher-level commands are not taught adequately, are difficult to implement, or are simply unnecessary for most users. Most of all, the results of error-behavior studies suggest the need for continual evaluation of systems so that the problems can be identified and the systems improved.

Future Research

Information systems have not yet reached the stage of being user friendly for most of their users. We now know enough to begin to characterize the problems; much more work is required to find solutions for them. We need both design guidelines to alleviate known problems and basic research to identify general principles of user behavior. The initial groundwork for a psychology of human-computer behavior has been laid and research methods exist to continue the work. A base of implemented systems, available to a variety of user populations, exists for study. With sufficient devotion to research, we may soon have a class of "user friendly retrieval systems."

REFERENCES

1. Meads, Jon A. "Friendly or Frivolous?" *Datamation* 31(1 April 1985):96-100.
2. *Webster's Ninth New Collegiate Dictionary*, s.v. "user."

3. ________ , s.v. "friendly."

4. Matthews, Joseph R., and Williams, Joan Frye. "The User Friendly Index: A New Tool." *Online* 8(no. 3, 1984):31-35.

5. Meads, "Friendly or Frivolous?"

6. Gentner, Dedre, and Stevens, Albert L., eds. *Mental Models*. Hillsdale, N.J.: Erlbaum, 1983.

7. Lewis, C., and Mack, R. "Learning to Use a Text Processing System: Evidence from 'Thinking Aloud Protocols.' " *Proceedings of the Human Factors in Computer Systems Conference* (Gaithersburg, Md., 15-17 March 1982). New York: Association for Computing Machinery, 1982.

8. Halasz, Frank. "Mental Models and Problem Solving Using a Calculator." Ph.D. diss., Stanford University, 1984.

9. Borgman, Christine L. "The User's Mental Model of an Information Retrieval System: Effects on Performance." Ph.D. diss., Stanford University, 1984; and ________ . "The User's Mental Model of an Information Retrieval System: An Experiment on a Prototype Online Catalog." *International Journal of Man-Machine Studies* 24(no. 1, 1986):47-64.

10. Dickson, Jean. "An Analysis of User Errors in Searching an Online Catalog." *Cataloging & Classification Quarterly* 4(no. 3, 1984):19-38.

11. Card, Stuart K., et al. *The Psychology of Human-Computer Interaction*. Hillsdale, N.J.: Erlbaum, 1983.

12. ________ . "Evaluation of Mouse, Rate-Controlled Isometric Joystick, Step Keys, and Text Keys for Text Selection on a CRT." *Ergonomics* 21(1978):601-13; and ________ , *The Psychology of Human-Computer Interaction*.

13. Egan, Dennis E., and Gomez, Louis M. "Assaying, Isolating and Accommodating Individual Differences in Learning a Complex Skill." In *Individual Differences in Cognition*, vol. 2, edited by R.F. Dillon. New York: Academic Press, 1985.

14. Sitton, S., and Chmelir, G. "The Intuitive Computer Programmer." *Datamation* 30(no. 16, 1984):137-40; and Lyons, M.L. "The DP Psyche." *Datamation* 31(no. 16, 1985):103-10.

15. Alspaugh, C.A. "Identification of Some Components of Computer Programming Aptitude." *Journal for Research in Mathematics Education* 3(no. 2, 1972):89-98; Butcher, D.F., and Muth, W.A. "Predicting Performance in an Introductory Computer Science Course." *Communications of the Association for Computing Machinery* 28(no. 3, 1985):263-68; Coombs, M.J., et al. "Learning a First Computer Language: Strategies for Making Sense." *International Journal of Man-Machine Studies* 16(no. 4, 1982):449-86; Fowler, G.C., and Glorfield, L.W. "Predicting Aptitude in Introductory Computing: A Classification Model." *Association for Educational Data Systems Journal* 14(no. 2, 1981):96-109; Stevens, D.J. "Cognitive Processes and Success of Students in Instructional Computer Courses." *Association for Educational Data Systems Journal* 16(no. 4, 1983):228-33; and Stevens, D.J., et al. "Group Differences in Computer Science Aptitude." *Association for Educational Data Systems Journal* 14(no. 2, 1981):84-95.

16. Bourne, Charles P. *Analysis of ERIC On-Line File Searching Procedures and Guidelines for Searching*. Berkeley: University of California, Institute of Library Research, 1974; Fenichel, Carol Hansen. "Online Information Retrieval: Identification of Measures that Discriminate Among Users with Different Levels and Types of Experience." Ph.D. diss., Drexel University, 1979; Fenichel, Carol Hansen. "Online Searching: Measures that Discriminate Among Users with Different Types of Experience." *JASIS* 32(no. 1, 1981):23-32; and Oldroyd, B.K., and Citroen, C.L. "Study of Strategies Used in On-Line Searching." *On-Line Review* 1(no. 4, 1977):293-310.

17. Fenichel, Carol Hansen. "The Process of Searching Online Bibliographic Databases: A Review of Research." *Library Research* 2(no. 2, 1980):107-27.

18. ________ , "Online Information Retrieval"; and ________ , "Online Searching."

19. Penniman, W. David. *Modeling and Evaluation of On-Line User Behavior* (final report to the National Library of Medicine). Dublin, Ohio: OCLC, 1981.

20. Brindle, Elizabeth A. "The Relationship Between Characteristics of Searchers and Their Behaviors While Using an Online Interactive Retrieval System." Ph.D. diss., Syracuse University, 1981.

21. Bellardo, Trudi. "Some Attributes of Online Search Intermediaries That Relate to Search Outcome." Ph.D. diss., Drexel University, 1984.

22. Woelfl, Nancy N. "Individual Differences in Online Search Behavior: The Effect of Learning Styles and Cognitive Abilities on Process and Outcome." Ph.D. diss., Case Western University, 1984.

23. Borgman, Christine L. "End User Behavior on an Online Information Retrieval System: A Computer Monitoring Study." *Association for Computing Machinery Special Interest Group on Information Retrieval. Forum* 17(no. 4, 1983):162-76; ________. *End User Behavior on the Ohio State University Libraries' Online Catalog: A Computer Monitoring Study* (Research Report OCLC/OPR/RR-83/7). Dublin, Ohio: OCLC, 1983; Larson, Ray R. *Users Look at Online Catalogs: Results of a National Survey of Users and Non-Users of Online Public Access Catalogs. Part 2: Interacting with Online Catalogs.* Berkeley: University of California, Division of Library Automation and Library Research and Analysis Group, 1983; and Tolle, John E. *Current Utilization of Online Catalogs: Transaction Log Analysis* (final report to the Council on Library Resources. Research Report OCLC/OPR/RR-83/2). Dublin, Ohio: OCLC, 1983.

24. Matthews, Joseph R., et al. *Using Online Catalogs: A Nationwide Survey.* New York: Neal-Schuman, 1983.

25. Borgman, Christine L. "The User's Mental Model: Effects"; and ________, "The User's Mental Model: Experiment."

26. ________. "Individual Differences in the Use of Information Retrieval Systems: A Pilot Study." *Proceedings of the 49th Annual Meeting of the American Society for Information Science* (28 Sept.-2 Oct. 1986), vol. 23, edited by Julie M. Hurd, pp. 20-31. Medford, N.J.: Learned Information, 1986.

27. Woelfl, "Individual Differences in Online Search Behavior."

28. Borgman, Christine L. "Why are Online Catalogs Hard to Use? Lessons Learned from Information Retrieval Studies." *JASIS* 37(no. 6, 1986):387-400.

29. ________. "Psychological Research in Human-Computer Interaction." In *Annual Review of Information Science and Technology*, edited by Martha E. Williams, pp. 33-64. White Plains, N.Y.: Knowledge Industry Publications, 1984.

30. Lancaster, F. Wilfrid, et al. "Evaluating the Effectiveness of an Online Natural Language Retrieval System." *Information Storage & Retrieval* 8(no. 5, 1972):223-45; and Sewell, W., and Bevan, A. "Nonmediated Use of MEDLINE and TOXLINE by Pathologists and Pharmacists." *Bulletin of the Medical Library Association* 64(no. 4, 1976):382-91.

31. Fenichel, "Online Information Retrieval"; and ________, "Online Searching."

32. Penniman, *Modeling and Evaluation of On-Line User Behavior.*

33. Tolle, John E., and Hah, Seh-chang. "Online Search Patterns: NLM CATLINE Database." *JASIS* 36(no. 2, 1985):82-93.

34. Tolle, John E. "Understanding Patrons' Use of Online Catalogs: Transaction Log Analysis of the Search Method." In *Proceedings of the 46th American Society for Information Science Annual Meeting*, vol. 20, edited by F. Vondran et al., pp. 167-71. White Plains, N.Y.: Knowledge Industry Publications, 1983; and ________, *Current Utilization of Online Catalogs.*

35. Borgman, "End User Behavior"; and ________, *End User Behavior on the Ohio State University Libraries' Online Catalog.*

36. Dickson, "An Analysis of User Errors."

37. Taylor, Arlene G. "Authority Files in Online Catalogs: An Investigation." *Cataloging and Classification Quarterly* 4(no. 3, 1984):1-17.

38. Ibid.

39. Fenichel, "The Process of Searching."

40. ________, "Online Information Retrieval"; ________, "Online Searching"; Martin, W.A. "A Comparative Study of Terminal User Techniques in Four European Countries on a Large Common On-Line Interactive Information Retrieval System." In *First European Congress on Documentation Systems and Networks* (held in Luxembourg, 16-18th May 1973), edited by the Commission of the European Communities, Directorate General, pp. 107-67. Luxembourg: Commission of the European Communities, 1973; and Oldroyd, and Citroen, "Studies of Strategies Used in On-Line Searching."

41. Wanger, Judith, et al. *Impact of On-Line Retrieval Services: A Survey of Users, 1974-1975*. Santa Monica, Calif.: Systems Development Corp., 1976.

42. Fenichel, "Online Information Retrieval"; and ________ , "Online Searching."

43. ________ , "The Process of Searching."

44. ________ , "Online Information Retrieval"; and ________ , "Online Searching."

45. Matthews, et al., *Using Online Catalogs*.

46. Ibid.

47. Markey, Karen. *Subject Searching in Library Catalogs*. Dublin, Ohio: OCLC, 1984.

48. Kern-Simirenko, C. "OPAC User Logs: Implications for Bibliographic Instruction." *Library Hi Tech* 1(no. 3, 1983):27-35.

49. Johnson, M.F. "An Analysis of the Log of an Online Public Access Catalog." St. Louis: Washington University, School of Medicine Library, 1982 (unpublished manuscript).

50. Dickson, "An Analysis of Errors."

51. Taylor, "Authority Files in Online Catalogs."

52. Borgman, Christine L., et al. "Incorporating User's Information Seeking Styles into the Design of an Information Retrieval Interface." *Proceedings of the 48th ASIS Annual Meeting*, vol. 22, edited by Carol A. Parkhurst, pp. 324-30. White Plains, N.Y.: Knowledge Industry Publications.

DALE K. CARRISON
Director
Minnesota State University System
Project for Automated Library Systems
Mankato State University
Mankato, Minnesota

Is "User Friendly" Really Possible in Library Automation?

Everyone wants to have a "user friendly" system, but no one can really define what user friendly is. A friend suggests that we should be talking about "user seductive" rather than user friendly. The person using an online catalog should perceive a natural relationship and not have to stop to think about his or her interaction with the computer. The truth of the matter is that an online catalog is so much better than a card catalog that library patrons will put up with a lot of "unfriendly" things in an automated system.

The term *user friendly* has become a buzzword. Everyone would probably agree that online library systems should be approachable. However, despite efforts to make system use easier, many first-time users still feel intimidated. The major cause of user fear may be the everyday jargon used by those persons who are the corporate keepers of the Holy Grail—i.e., automated library systems. Often the words used in discussing online systems are overly expressive and needlessly violent in tone. Even the term *user* fits this situation because it sounds drug related rather than library related.

Another example is one that probably causes Reverend Jerry Falwell consternation—i.e., "abort, retry, or ignore." Why not use "quit, try again, or bypass" instead? Or another example is a multitasking communications program that goes into a "deadly embrace" and so informs the person at the terminal.

The key is to simplify dialogue and use everyday terminology instead of jargon. It is just as easy to say "program error" as to say "bug." "Running" a program sounds better than "executing" it. And a "system failure," no matter how unwelcome, is not nearly as intimidating as a "crash" which sounds positively life threatening. Where does the computer world get so much emotionalism?

Other examples of misuse of the language are prevalent in the computer world. The response should be to promote proper usage and decry those who use meaningless or ambiguous words. A need exists to step back and not take ourselves too seriously. If this process does not occur, a large segment of library patrons will be resistant to using automated library systems effectively and efficiently.

Libraries with automated library systems have found that patrons naturally gravitate to the terminals of the online catalog because of their belief that a "computerized" catalog provides better access to the library materials that they want to use. This situation is important in a society that places value on innovation, speed, and ease of use for all services whether they be in libraries or other places. User sophistication is constantly growing through increased access to automated library systems. Expectations of ease of use are also increasing.

The online catalog should assist experienced and inexperienced users by revealing its inner working organization. A person should be able to use an online catalog without knowing its structure and without regard to age, education, experience, or sex. Minimal instrution should be required (less than fifteen minutes).

User friendliness is based on positive interaction between the computer and the terminal user. Human conversation provides the best model for this human-computer communication. The problem-solving dialogue component of conversation is particularly applicable to human-computer applications. In the problem-solving, task-sharing activities of online catalog use, something close to symbiosis between the person and the computer must be attained. A close union and a result-oriented interdependence must be established. Several properties of dialogue make it a good form of communication for human-computer activity. Stewart, in his article "Communicating with Dialogues," reminds us that: "A dialogue is by definition a two-way process....[It] involves the sharing of knowledge by the exchange of information." Successful dialogues according to Stewart have five requirements:

1. Both parties must be able to send and receive information without undue constraint.
2. The language chosen as the carrier of the dialogue must facilitate the expression of subtle or complex ideas.
3. Accordingly, the interpretation and understanding of the language—its symbols and rules—must be shared to the same degree by both parties.
4. Each party must be able to grasp the context of meaning of the other party and successfully follow or "keep up" when that context shifts or changes.

5. This, in turn, requires being responsive to the feedback of the other party and "modifying the communication so that it is better suited to the other party."[1]

The literature contains much information about the general requirements for effective human-computer communication. Perhaps the experience with online catalogs is best explained by the observations of Hayes, et al. in an article "Breaking the Man-Machine Communication Barrier." He states:

> Simply put, today's systems are not very good at communicating with their users. They often fail to understand what their users want them to do and then are unable to explain the nature of the misunderstanding to the user. In fact, it is the common experience of users of interactive systems, whether novice or experienced, infrequent user or regular, that communication with their machines is a time consuming and frustrating experience.

Why does this barrier exist? According to Hayes, the computer "lacks the basic communication skills that come so easily to almost all of us."[2] Mutual understanding is a necessity in good dialogue. Partners in a dialogue must be ready to repeat, simplify, or otherwise clarify the message. The timeliness of message acknowledgment and response is also important for effective dialogue as explained by Hildreth in his OCLC Technical Report "Optimal Response Times in an Online Interactive Computing Environment."[3]

Since language is the medium of message transfer, every effort must be made to use semantics and syntax that facilitate the dialogue between the computer and the user in terms of vocabulary, sequence, and consequences. The following things are important when considering semantics and syntax:

Semantics

1. Completeness: the set of commands provided must be capable of evoking all the functions required in the user's model of the task domain.
2. Functionality: the operational interrelationships among the commands must be flexible enough to be compatible with a variety of user activity patterns in online information retrieval.
3. Singularity: each function should be represented by only a single, distinct command, and no command should be capable of evoking multiple functions.
4. Modularity: a basic, simple subset of the language, capable of evoking the fundamental functions, should be provided for the new or occasional user. The design structure must be hierarchical, permitting the learning and use of specialized extensions to the basic subset to develop in a linear manner.

5. Variety: extensions to the basic command core representing specialized and more sophisticated features should be provided for the experienced user.
6. Linearity: the structure of the basic command core and its extensions should permit the user to proceed in a left to right, top down, manner.
7. Optionality: alternative ways of expressing commands during input should be supported consistent with the syntax criteria listed later. This would include the use of abbreviations, flexible argument order, and user designation of default values for commands and parameters.
8. Simplicity: any complexity introduced should result only from the extension of the basic command core to include more powerful or refined capabilities and not from the meaning of the commands or the syntax governing their use.
9. Consistency: command words should have the same meaning in all contexts, and syntax decisions should be applied uniformly throughout the language.

Syntax

1. The structure of the command statements should be patterned on familiar, natural language phrase structures.
2. Command keywords should be short, familiar words chosen from the user's natural language and should clearly express the specific action being commanded.
3. Self-evident abbreviations should be permitted but governed by easy to remember abbreviation rules.
4. Punctuation required for command construction should be kept to a minimum and be limited to universally familiar symbols. Blank spaces should serve as delimiters.
5. The positioning of arguments following a command word should be flexible and not governed by a predefined, rigid ordering of components.
6. Command syntax should be compatible with display syntax.
7. Default states for commands and parameters, as well as the capability to set or reset them, should be provided for the user.
8. The user should not have to learn specialized keyboard techniques. The use of special function keys or control key combinations should be kept to a minimum. Entry in either upper or lowercase should be permitted.

During the past five years several research studies related to online catalogs have been completed. Most of these studies are useful in providing insight into patron attitudes about online catalogs. Virtually all of this research has shown that both users and nonusers have a positive feeling about online catalogs. Unfortunately, these same studies do not provide a clear picture of what really constitutes a user friendly online catalog or even what causes a user to express satisfaction and be positive. No research to date has conclusively identified online system features required to "optimize" the human-computer interface.

As Hildreth states in his book *Online Public Access Catalogs: The User Interface*, online catalogs are experiencing a period of intense development and deployment. Consequently, the emphasis should be on inclusion of interface features that are generally accepted as user friendly through available evidence, if not actual consensus among designers and experienced users. Hildreth sums up his comments by stating: "The need now is to *improve* the interface, not optimize it once and for all time."[4] From a technical point of view, online catalogs are at a stage when software development techniques must catch up and keep pace with what is known about good dialogue between persons and computers in particular and user-oriented interfaces in general. Gaines, in his article "The Technology of Interaction," reports that there is not yet agreement among software designers as to what constitutes good dialogue programming. He states: "We are in a position today in programming man-computer interaction that we were with hardware design thirty years ago and software design ten years ago."[5]

In addition to the human-computer communication, environmental considerations also affect the ease of use of an online catalog. Consideration should be given to location of terminals, tables used for terminals, and the physical characteristics of terminal design. However, experience with online library automation systems indicates that color display, graphics, touch screen, and monitor display color (amber, green, or monochrome) are of minor concern. Other important considerations are level and completeness of documentation, display of diacritics, and invisibility of hardware. The provision of comprehensive "help" modules, which are easily accessible and contact sensitive, is important to a user friendly environment. This "help" should include user requested help, system defined help, automatic response prompts, and informational messages.

In the area of user friendliness for handicapped library patrons, the University of California's Division of Library Automation has developed a "talking terminal" that enables blind and visually impaired users to use the MELVYL Online Catalog without assistance. The terminal unit incorporates a keyboard labeled with large letters and braille, a screen display that magnifies the image, and a digital voice synthesizer. The terminal responds to users by speaking, by providing a tape from a built-in cassette tape recorder, and by providing a display up to sixteen times its normal size. One drawback is the current cost—which is $10,000—plus the cost of a braille printer if needed.

Another approach to user friendliness is to give terminal users the opportunity to communicate questions, comments, and observations. The LIAS system at Pennsylvania State University uses the OPPS command, which allows any user to enter a free-form message. These messages are periodically printed out for analysis and for appropriate reaction or response. This command can be easily used without any intervention of a staff member. One limitation is that it lacks a means of direct response to the person who initiated the command. Thus there is no way to respond to a request for information.

Based on observation and discussion with users of online catalogs, the following criteria must be considered in developing a user friendly online catalog interaction with the user:

1. Be conversational but instructional at the same time.
2. Use consistent format and terminology.
3. Use mnemonic devices whenever possible.
4. Use formats which facilitate understanding of information presented on screen.
5. Eliminate as many steps as possible.
6. Be positive in response statements.
7. Be forgiving of errors in entry.
8. Accommodate many user punctuation and spacing inputs.
9. Provide computer response time averaging three seconds or less.

The current emphasis by the whole library automation industry is to make systems more user friendly. This focus is underlying all current marketing and design efforts, whether in the area of hardware—such as terminals—or in the area of the software which runs the system. User friendliness is an ever present force in library automation today. Within the next five years, we will wonder why we ever did things in the old laborious way—i.e., keying in commands. This process will all be superseded by voice commands or by a mixture of voice and graphics. We are still some years away from one person's dreams—a truly portable terminal which can be taken out jogging.

REFERENCES

1. Stewart, T.F.M. "Communicating With Dialogues." *Ergonomics* 23(March 1980):909-19.

2. Hayes, P., et al. "Breaking the Man-Machine Communication Barrier." *Computer*, March 1981, pp. 19-30.

3. Hildreth, Charles R. *Optimal Response Times in an Online Interactive Computing Environment*. Dublin, Ohio: OCLC, 1981.

4. ———— . *Online Public Access Catalogs: The User Interface*. Dublin, Ohio: OCLC, 1982.

5. Gaines, Brian R. *The Art of Computer Conversation*. Englewood Cliffs, N.J.: Prentice Hall, 1984.

EMILY GALLUP FAYEN
Assistant Director of Libraries
Van Pelt Library
University of Pennsylvania

User Interfaces for Online Library Catalogs

As computer systems moved out of the laboratory, it became apparent that many of them were intended to be used by people who had no knowledge of programming or understanding of computer systems. It soon became necessary to develop more congenial ways for people to interact with systems, and the whole notion of the "computer-user interface" or the "man-machine dialogue" was born.

In general, the complexity of a computer system can be categorized as shown in figure 1. The figure illustrates how systems complexity increases depending on the number of programs or modules it contains and who will use it. The simplest case, of course, is a single program written by one person for his or her own use. System complexity is increased by two major factors: (1) if the system includes more than one program which must work together, and (2) if the system is to be used by someone other than the programmer who wrote it. Obviously, programmers can get away with leaving an error message that says something cryptic like "zero-length response not allowed" or worse, but if a computer system is to be used by others, the instructions for using the system and any error messages that may be necessary must be clear to the user.

Online library systems are a classic example of complex systems of programs designed for use by others. Library systems are made up of many intricate programs with complex relations among them. The library database consists of numerous files which are interrelated in various ways. Library users—both staff and patrons—need to perform a large number of complex actions using these programs and files. The challenge to the library systems designer is to create an effective user interface for this environment.

Early attempts at developing effective interfaces assumed that it was a relatively trivial matter to design a few menus and other screen displays to guide the user through the system and provide assistance when needed.

```
-------------------------------------------------------
|      S Y S T E M S   C O M P L E X I T Y ------>    |
|                          |                          |
| S                        |                          |
| Y                        |                          |
| S |   ONE PROGRAM        |   ONE PROGRAM            |
| T |   FOR USE BY         |   FOR USE BY             |
| E |   PROGRAMMER         |   OTHERS                 |
| M |                      |                          |
| S |                      |                          |
|___|______________________|__________________________|
| C |                      |                          |
| O |                      |                          |
| M |    MULTIPLE PROGRAM  |   MULTIPLE PROGRAM       |
| P |    SYSTEM FOR USE    |   SYSTEM FOR USE         |
| L |    BY PROGRAMMER     |   BY OTHERS              |
| E \/                     |                          |
| X                        |                          |
| I                        |                          |
| T                        |                          |
| Y                        |                          |
-------------------------------------------------------
```

Figure 1. How Computer Systems Become Complex

Many bad experiences have shown that designing effective user interfaces is neither simple nor easy. The process of developing a new user interface may appear to be simple because the best interfaces are simple and easy to use. But this very simplicity is deceptive. It usually means that under the surface, some very sophisticated language parsers, menu displays, and command syntax structures are in place that effectively hide all the system's complexities from the user.

Another difficulty facing designers of user interfaces is that they are usually in the position of creating a so-called "user friendly front end" or interface to an existing system. Although cleverly constructed menus and screen displays can help to make a system more user friendly, some of the problems that arise may be caused by a poorly designed underlying system. These difficulties may be almost insurmountable. For example, some online systems use a derived search key (e.g., first four letters of the author's name, followed by a comma, followed by the first three characters of the first significant word of the title). Explaining how to construct this search key to the novice user in a friendly fashion is virtually impossible. The best solution may be to develop a parser that analyzes the user's search request and constructs the search key. This approach may not be entirely satisfactory because the search results produced by the derived key may not be at all what the user wanted.

Perhaps in the future some new library systems designer will think about how users ought to be able to get information out of an online library system and then will create a system to support it. Apple took this approach with the Macintosh personal computer. Computer designers at Apple had many new ideas about how users ought to be able to work with a personal computer and they then set about developing a machine to allow these new approaches. Cambridge Library System's original touch screen interface to the library online catalog was an early attempt to provide such a new mode of access. As is well known now, this effort was partially successful. It was well liked by users in small libraries but found to be too slow and too cumbersome by patrons in large libraries. Nevertheless, it was a brave attempt at finding an entirely new approach to online library catalog access. Similarly innovative efforts must be encouraged in the future.

Features of a User Friendly System

A user friendly system has a few well-known characteristics. These may be implemented in quite different ways, but the results as far as the user is concerned will be quite similar. The main features are:

—*Users cannot get "lost" in the system.* The system always lets the user know what is being done and offers suggestions about what to do next. There are no surprises. The system does not, at the touch of a finger, go off and perform some action that is totally incomprehensible to the user.

—*Users cannot enter illegal commands.* The system prompts for certain user actions. If the anticipated choice is not made, the system does nothing or prompts again for the expected data perhaps at the same time suggesting an appropriate choice.
—*Users can choose among several different simple modes of data entry.* One keystroke or perhaps pressing the enter key is all that is required of the user. Some systems that are equipped with a mouse or other pointing device allow users to enter data using the device or using the keyboard.
—*Users are in control.* The user is given enough information about the system to understand how it works and how to make it do what the user wants. The user has a mental model of the system.

User Interfaces for Library Applications

Online library systems need good user interfaces to be successful. Although the basic concepts of online information retrieval are relatively simple, the features of many online catalog systems are not simple. Further, the library card catalog itself is not simple or easy to use. Thus an effective user interface for an online library system has to deal with complex record structures, data types, and files and it must be a comfortable tool for the occasional library user as well as the experienced staff member. This is asking a lot. The specifications for most library user interfaces have a built-in multiplicity of functions that is seldom questioned. For example:

—Is the user interface intended to teach users how to use the online catalog or how to use the library or both?
—Is the user interface intended for use by library patrons or staff or both?
—Is the user interface intended to make the system easy to use or easy to learn or both?
—Is the user interface intended to serve the infrequent or casual user, the expert user, or both?

If the answer is "all of the above" it is no wonder that it has proven difficult to develop user interfaces for online library systems that are adequate.

Modifications to the NOTIS Screen Displays for PennLIN

The Northwestern Online Total Information System (NOTIS) software is being installed at the University of Pennsylvania to form the heart of PennLIN, Penn's Library Information Network. As part of this process, the NOTIS screen displays are being modified to suit the Penn environment.

NOTIS is a command-driven system; that is, the underlying design philosophy in constructing the system is that users should be able to enter any command for which they are authorized at any point in the session. Therefore in NOTIS, the screen displays that have been developed for Northwestern's Library User Information System (LUIS) are really prompts for commands rather than true menus. One of the difficulties that arises is how to present all possible options to users without overwhelming them with choices. Another problem is how to move from the command prompt screens to help screens and back without confusing users. Figure 2 shows the NOTIS screen displays and the linkages between them.

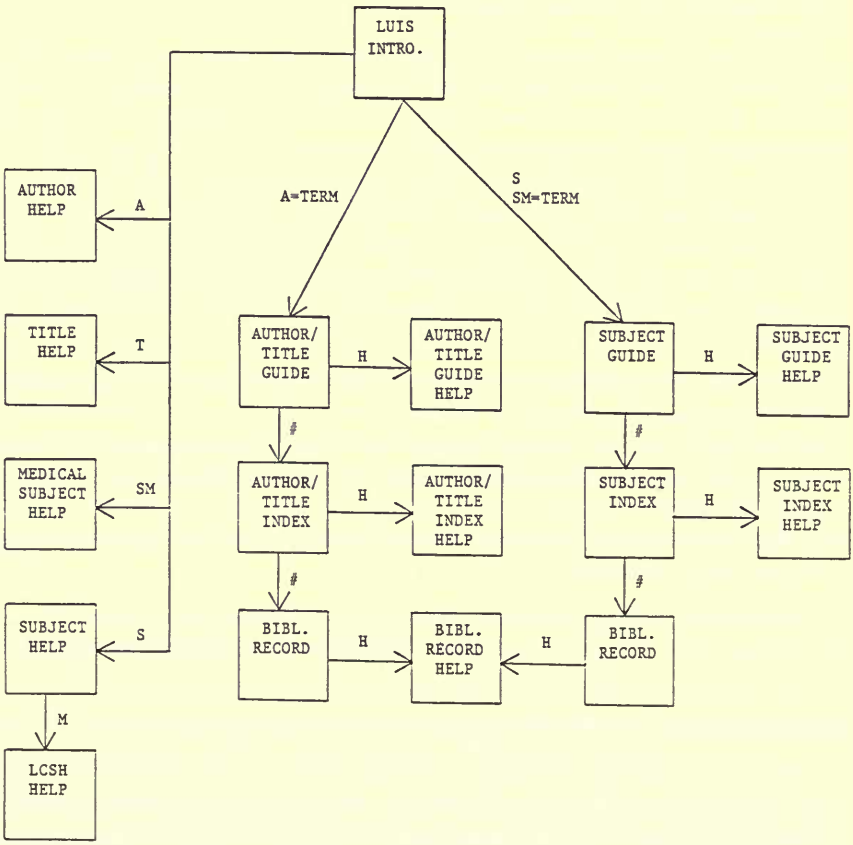

Figure 2. NOTIS/LUIS Screen Displays

Although the choice of wording on the screen displays is very important in creating a user friendly system, having an underlying structure for the screens and menus that is clear to the user is even more important. It is this structure that serves to:

—keep users from getting "lost" in the system;
—keep users in control; and
—give users a mental model of the system.

In a recent article, Ben Schneiderman has stated that "the primary task for menu designers is to create a sensible, comprehensible, memorable, and schematic organization."[1] He likens this organization to a breakdown of a larger whole into its parts such as the chapters in a book, a catalog into sections, or a restaurant menu into categories such as appetizers, soups, and desserts. Schneiderman further categorizes menu systems into two main classes: (1) single-menu systems—that is, a system with one menu offering binary or multiple choices; and (2) sequences of menus. These may be linear, tree structured, or cyclic or acyclic networks. Simple menu systems cause few difficulties for users because they offer limited options, but more complex menu structures require that users understand how to move around the system—that is, how to navigate the sequences of menus.

Very little research work has been done on the best format and content of screen displays. Empirical findings have led to some conclusions, but much more work needs to be done. Preliminary findings indicate that graphic design and layout of the displays are extremely important, but there are almost no research studies that have focused on library applications. Opinion varies on the use of color to enhance computer displays for patrons or staff, but no definitive research results are known which support the superiority of color or monochrome displays for use by patrons or staff.

Some conclusions gathered from observation and experience are:

—don't use jargon;
—especially don't use computer jargon;
—use familiar and consistent terminology;
—use consistent and concise phrasing;
—make screens consistent in format and terminology from frame to frame;
—use the same area of the screen to display prompts or for user data entry from screen to screen;
—don't make the screens too full;
—don't overwhelm the user with choices; five to seven options are optimal;
—make sure that menu choices are clear and do not overlap;
—try to eliminate error messages. Users should not be able to do anything "wrong";

—give users a hint about the result of menu choices—that is, there should be no surprises;
—use letters or numbers for menu choices but not both;
—letters are preferred for menu choices because of their mnemonic value, but letters must be selected with care;
—put the most important part of the instruction at the beginning of the line;
—keep the user in control. Have an "undo" key if possible; and
—keep it simple, remembering that what seems simple to the user may mean that very sophisticated processing is taking place in the background.

Research Topics for Developing User Friendly Systems

As indicated earlier, there are many topics in the design of user interfaces that are poorly understood. There is a great opportunity for testing the online library systems now in use to see what works and, conversely, what does not and to test new ways of handling various interface problems. Some unanswered questions that need research follow:

—When specifying a menu choice, should the action precede or follow the explanation in the display—e.g.,

Type A To search for an author

or

To search for an author Type A

—What area of the screen display should be reserved for user data entry?
—What is the best place on the screen to use for error messages (if any)?
—Highlighting seems to be effective if not overused. How is it best employed?
—What about color displays? Is color just a frill? Does it have any use besides its value as a novelty?
—What indicators on the screen (e.g., leader dots, dashes, arrows, blank space) work best to connect menu choices with the appropriate actions?
—How is white space best used?
—What is the best way to handle a menu or screen display that is too big to fit on one screen?

Specifically for the library online catalog, a few other areas in which research results would be extremely useful:

—*In what order should bibliographic records appear? Should it be the same for staff and for public users*? Some online library systems display bibliographic citations in conventional author-title order while others display the most recent additions to the file first. What order is preferred? Should users have a choice? Should users learn how to do online sorting of records?

—*What fields or data should appear for each bibliographic record? Should there be a default which users may change?* Research has shown that patrons use very little of the information that is available on the conventional library catalog card. Are all of the data elements necessary? Which ones should users see? Should users be able to select different levels of completeness?

—*What techniques work best to help users whose search strategy nets too few results or too many?* Preliminary analysis of online catalog transaction logs shows that users whose search strategies retrieve too many or too few citations have difficulty in narrowing or broadening the search. Is this a fault of the system? Is it because users don't understand retrieval of sets? A related issue is how much information is enough information and what is the library's responsibility to make sure users get all available information whether they want it or not?

—*How can one determine if a search that results in no hits represents a failure of the system or a failure of the database?* Several studies have shown that searches of online catalogs sometimes do not retrieve the desired records when the records are in fact in the database. Sometimes it is difficult to determine if this is the fault of the system or the fault of the search strategy. In other cases, the desired records really are not in the database. What tests can be made to increase confidence in the retrieval effectiveness of online systems?

—*What is the best way to explain the basic idea of retrieval of sets to library users?* Analyses of online catalog transaction logs and user comments reveal that some patrons and staff members really do not understand the concept of retrieval of sets of citations from an online database. Some of these same users have had a formal introduction to set theory, but apparently there is no carry-over to the library application. How can the interface give these users a good mental model of the online database so that they will truly understand how the system works? Should online catalogs emulate the card catalog? Are there other models that are more appropriate?

Conclusion

User interfaces have come a long way since the earliest attempts to design user friendly systems. Few of today's systems resort to the withering "illegal command" or other decidedly unfriendly responses. Most of the new systems are no longer user hostile or user vindictive, although some are still rather opaque and require that users learn more about the system than most of them ever wanted to know. Still, thousands of happy users use hundreds of these systems every day, but there is still a long, long way to go before online systems are truly user friendly.

In the future, advances in artificial intelligence and expert systems may offer a path to creating user interfaces that are more natural; that is, they will allow queries to be posed in natural language and help users to select the most appropriate sources for searching. But in the near term, if systems designers applied the empirical results gleaned from online catalog use and other studies that are available already, online systems would improve. More information about how users would like systems to operate and a much better understanding of the fine art of designing screens and constructing menus will help to make online systems much more acceptable to users. Perhaps that is what we really mean by user friendly.

REFERENCE

1. Schneiderman, Ben. "Designing Menu Selection Systems." *JASIS* 37(no. 2, 1986):57.

GARY A. GOLDEN

Director
Camden Library
Rutgers, The State University of New Jersey

Taming the Unfriendly System: Microcomputers as Patron Terminals to Access an Online Catalog

Two concepts capture the spirit of this discussion. The first is a quote by Marshall McLuhan who said about technology: "If it works it's obsolete."[1] The other is Finagle's Law Of Information which states:

> The information you have is not what you want
> The information you want is not what you need
> The information you need is not available.[2]

Automation in libraries has been on a rapidly moving roller coaster over the past decade. At first the major concerns were whether to automate with the few existing vendor systems. Many libraries designed their own systems while others shopped around or waited. Over the past five years many turnkey systems promising an automation nirvana for libraries were developed. This conference and the growing body of literature shift this emphasis toward making these systems easier to use. An even newer concern is how to develop integrated systems or "information gateways" to allow access to expanding internal and external databases.

The rallying cry for librarians is not whether to automate, but can we automate our many internal processes and access external systems from the same terminal. Automation, however, takes place within the constraints of limited money. The question becomes, How can we allow all patrons easy and quick access to a new world of information choices?

We have slowly—and sometimes quite painfully (in terms of money and staff morale)—come to the realization that automation might not be the panacea. The programming proverb "Garbage in—garbage out" is haunting many a library committee as its members wrestle with past cataloging and classification inconsistencies. Automating our catalogs or circulation functions means users must develop entirely new habits in their quest for information. In addition, making only our catalogs available online is not enough for many patrons. On many occasions we have found patrons attempting to locate citations to journal articles in our online catalog. When told they couldn't do this, their most common responses were Why? or I thought that computers could do this.

Searching these online catalogs is another problem area. In many searches more than one patron-initiated action is necessary to find a citation to a specific book (e.g., a title search finding many matches might need narrowing). Remembering a long set of commands or different search strategies for authors, titles, or subjects is difficult for patrons who are infrequent library users. Having thousands of new freshmen, graduate students, and faculty needing assistance each year can strain an already difficult bibliographic instruction program. The cost in staff morale is immeasurable.

In user friendly systems, symbols are explained and patrons are led through this maze of commands. Many librarians need help with systems that are not friendly or friendly enough. How can they improve these systems? Is there a system flexible enough to allow access to numerous systems at the same time? The answer to all these questions is microcomputers. Before proving this to you, I would like to present an overview on the direction of catalog and circulation automation within Illinois academic libraries.

Automation in Illinois

At present there is an automated short-record circulation system called LCS. LCS is the largest cooperative resource sharing network located within a single state. Every academic library desiring state money for automation must join LCS. There are twenty-seven member libraries including every four-year state-supported university and ten private institutions like DePaul and Judson College. As of April 1986, they have input 8.9 million titles and 15.4 million volumes into an IBM 3801 computer located in Chicago.

Members agree to share any circulating item with patrons associated with other member institutions. Delivery of materials takes seven to ten days and is through the Inter-System Library Delivery Service, an Illinois State Library funded operation. Each library has its own separate database, but every one of the 637 terminals can search each database individually. In FY86 there will be over 35 million searches on the system and over 300,000 interlibrary circulation transactions. Since the network's inception a little over five years ago, LCS schools have borrowed from each other over 1.25 million items.[3]

At the University of Illinois at Urbana-Champaign (UIUC), LCS is a known-item circulation system. The system came from Ohio State University (OSU) in the late 1970s and is totally different from the OSU system in record appearance and searching capabilities. Library locations, call numbers, and circulation status are given for every book or journal ever cataloged on the UIUC campus. In addition, serial holdings and order records are also included in LCS. Searches by author, author-title, title, specific call number, and a general call number range are possible.

A searching algorithm consists of a three-letter command followed by a slash and then a six-three or four-five combination of letters from the author's name or item's title. Books written by Kate Turabian are found in LCS by entering AUT/TURABIKAT followed by a carriage return. There are six different three-letter commands and four two-letter commands used to manipulate individual screens. In all there are ninety-six permutations of searches using these commands. In addition, there are 107 stopwords to remember not to use if they appear anywhere in a title or author's name. After getting a record the patron must know the symbols for locations and there are over 120 of these symbols. Searching another LCS school, patrons must put a two-letter code at the end of the command and then search the twenty-six databases separately. That could involve keying in sixteen characters twenty-six times (e.g., to search for Turabian at DePaul one does AUT/TURABIKAT/DP).

To make things more confusing, we have in the past year brought up our Full Bibliographic Record system or FBR for short. FBR contains materials cataloged in UIUC since we joined OCLC in late 1974. There are almost 1 million bibliographic records and over 2.3 million authority records. Included in the authority file are the Library of Congress name and subject heading tapes. A government grant to develop FBR stipulated that we include a public library in the development stages. The River Bend Library System agreed to test FBR with us. At present no LCS school participates in FBR.

FBR is searchable by subject, author, series, ISBN, ISSN, and by keyword title or keyword corporate authors. Boolean searches and right-side truncation are also possible for most searches. However, searches of FBR use a different command language and searching algorithm from LCS. There are over fifty possible searching algorithms using FBR. The stopword list used by FBR is smaller than that used by LCS. In calendar year 1985, over 5.5 million searches took place on FBR. After a FBR search, a link to the corresponding LCS circulation record must be made to find locations. We wanted to direct all known-item searches to LCS first because LCS contains records for everything ever cataloged at UIUC. There are 3 million titles in LCS *v.* only 1 million titles in FBR.

By the end of 1986, an ILLINET Statewide Online Catalog including everything cataloged via OCLC by every library in Illinois will exist. This catalog will contain over 3.4 million unique titles from more than 280 libraries. Using FBR searching algorithms, a user at any terminal will first search the local database and then will be able to look for an item anywhere in the state. Until FBR has an interlibrary loan module, only items from LCS schools will have a link to a corresponding circulation record. Other items will be available through the old interlibrary loan method.

Figure 1 shows some of the questions patrons must answer and understand before they can successfully search our system. Wrong answers to any question or not understanding the concepts behind these questions led to high failure rates in our online catalog. We have two command-driven systems that use different searching algorithms with different stopwords and access points. This also increased the failure rate.

Every time someone wanted to search for a book or journal they had to know what they were doing or else have help from a staff member. Printed search aids and online help screens are available at every terminal. However, the longer a patron sits at any terminal the less chance that someone else can do a search. The longer a search takes the more frustrated a waiting patron becomes and the longer the queues. In addition, our suggestion box was filled with complaints about our system. Since we have only eighty-nine public terminals we needed to be sure that patrons could search our online catalog rapidly, yet successfully. We also had to relieve the burden placed upon the staff of constantly explaining every search in two different systems. Staff morale was at an all-time low and sinking fast. We needed to make a totally unfriendly system user friendly.

Defining User Friendly

Before developing or accepting an interface to any system one must define the goals of that interface. We had eight criteria that we wanted our interface to meet. They were as follows:

1. that any patron can easily and rapidly use the system to find most things looked for without having to use commands or ask for help;
2. that patrons, in order to realize the full power of our system (i.e., being able to do every type of search), should be allowed to use commands from any terminal if they want;
3. that certain types of searches are not used that much and should be done only through the command syntax (e.g., a Boolean search of an ISBN and subject);
4. that the system response rate, which was less than a second in most searches and not more than three to four seconds in others, cannot be compromised;

```
1. Which system do I search?
   LCS for known item or location/circulation information
   FBR for subject, keyword, new material
2. What are the commands?   How are they entered?
   LCS -- AUT/ 6 - 3
   FBR -- F A   Last name  first name  middle name or initial
          T A   To search the authority file for correct
                heading
3. What does the three letter location code on LCS mean?
   STX = stacks  UGX = undergraduate library
4. How do I find call numbers/locations after searching FBR?
   Use link command
5. What if my subject heading does not work?
   Check online authority file
   Then issue find command to get bibliographic records
   Then use link command to get LCS record
7. How do I pick out one match from many on the screen?
   Use DSL/line number command
```

Figure 1. Thought Processes Necessary to Search Our Online Catalog

5. that the interface should improve some of the shortcomings inherent in any system;
6. that the interface with minor changes for location and circulation codes should be usable in all of the other LCS libraries;
7. that the interface's development be evolutionary and easy to change and update as the capabilities of the system changed; and
8. that the interface will not deter but assist us to incorporate future technological and software innovations into our system.

With these assumptions in mind, we decided there were two ways to implement a user friendly interface. One method involved the development of programs residing on the mainframe computer in Chicago. However, this was not viable because of two important reasons. First, it would increase the amount of communication between the mainframe computer in Chicago and hundreds of terminals around the state. This would not be an efficient use of either the communication lines or the mainframe. On a computer doing 35 million searches, response time could be degraded. The second reason involves the political nature of any large network. Change happens very slowly if at all. Program changes on the mainframe are difficult and time consuming because in this network, decisions for software updates are made by committee. The best solution to our problem of the unfriendly system was an interface developed and controlled at the local level.

An interface acts as a translator. The human speaks one language, the computer processor another, and the interface software mediates between the two. It transforms the general wishes of man into the exact commands demanded by the processor. We wanted these wishes to be correctly conveyed before they reached the mainframe. Employing intelligent terminals to access the mainframe was the answer. C.C. Cheng, a professor in our Linguistics Department, helped us toward this solution by developing and writing the interface.

The interface resides in IBM personal computers. The programming for the interface is primarily in BASIC with a small part in assembly language. These personal computers do not have disk drives and at present only search the online catalog. They were purchased in 1983 and at the time were much cheaper than models having disk drives. The program transfers from a floppy disk to a cassette and then the cassette runs the program into the personal computers. The transfer time is approximately four minutes per terminal. The personal computers stay on twenty-four hours a day. The interface needs reloading only when a change is made in the program or the power fails.

Cheng's premise behind the design of the interface was to program "a user interface aimed at capturing the natural processes of the user search and at providing a graceful interaction between the patron and the computer."[4] He wanted to unlock the power of the online catalog but at the same time have the patron need only typing skills to use the system. For those patrons who learn how to use our online catalog, the personal computers are also searchable using the command language. The following examples detail the power inherent in using the interface.

The interface is a menu with only five choices for searching (see fig. 2). All searches, except subjects, go first to LCS. If the search is unsuccessful, the patron could then search FBR. If the patron pressed the question mark key, he or she would go through one short help screen. The help screen has paragraphs comparing LCS with FBR, information about correcting mistakes, location of the return key, and information as to how to contact Cheng with suggestions. Since the interface is self-explanatory, this was the only help screen necessary to use the online catalog. The numbers in the upper right-hand corner are a real-time clock. This clock assists students in getting to class on time and may help to speed up users' time at the terminal.

```
INTERFACE READY ( SEE THE BOTTOM LINE TO SEARCH; PRESS <?>
         FOR HELP. )




PRESS 1 FOR AUTHOR-TITLE 2: TITLE 3: AUTHOR 4: CALL NO. ETC.
         5: SUBJECT
```

Figure 2. First Screen of the Interface

Figure 3 presents a typical title-search sequence. What the patron enters is shown in quotation marks for illustrative purposes. A normal search of our online catalog does not need quotation marks. What the interface actually sends to the mainframe is in small letters. This does not actually appear on the screen during a search. The patron answers yes if looking for a periodical. LCS has the ability to limit search results to serials only. The patron then is asked to give the first and second words of the title. If any of the 107 stopwords are input, the computer beeps and instructs the user to replace that word with another word. The interface takes the words input by the patron and puts them into the appropriate LCS searching algorithm (e.g., TLS/PETEPRINC). This command goes to the mainframe computer. Two matches return and the patron is asked for a line number (e.g., "2") and the interface does the LCS command DSL/2. An explanation of what appears on the screen can be seen by asking for help. A patron can then charge out the book. We allow patrons to charge their own books from every library on and off campus. Patrons can also search other libraries, and the interface then automatically searches the databases of all the other twenty-seven LCS schools. The searcher does not have to remember the two-letter codes or input the sixteen characters. The interface instructs the user as how to charge the book from another campus. Patrons with invalid identification cards (IDs) or holds on their IDs are instructed to ask for help. This whole process is rapid and mistake proof.

```
                    BEGINNING OF SEARCH

IS THIS TITLE A PERIODICAL? (PRESS <Y> IF YES <ENTER>
                                              OTHERWISE) no
TYPE THE FIRST IMPORTANT WORD OF THE TITLE AND
                                      PRESS <ENTER>: "PETER"
TYPE THE SECOND IMPORTANT WORD OF THE TITLE
                                    AND <ENTER>: "PRINCIPLE"
(IF NONE, PRESS <ENTER>; IF UNKNOWN, PRESS <-> AND <ENTER>)
TITLE SEARCH:        PETER PRINCIPLE

tls/peteprinc

SEARCHING CIRCULATION RECORDS

FOUND 2 SIMILAR ITEMS

LOOKING FOR EXACT MATCH.  ONE MOMENT . . .
```

Figure 3. Sample Title Search (part 1)

A search for books and journals on a specific subject begins in FBR (see fig. 4). The patron starts by typing a term or terms and can then limit the search if necessary (e.g., TOPIC—T, etc.). The interface puts this request into the proper search algorithm and searches the authority file

```
01 PETER, LAURENCE JOHNSTON, 1919-THE PETER PRINCIPLE$NY 1970
02 PETER, LAURENCE JOHNSTON, 1919-THE PETER PRINCIPLE$NY 1969
THESE ITEMS MATCH YOUR REQUEST
TYPE A LINE NUMBER TO SEE CALL NUMBER OF THAT ITEM AND TO
     CHARGE, RENEW, ETC.; OR PRESS < ENTER > TO GO ON

   "2"
dsl/2
658P441P PETER, LAURENCE JOHNSTON,1919-THE PETER PRINCIPLE$NY
    922826     1969     2  ADDED:  780702
01     001    16-4W    STX     RNEWD  860212/860427     UC
02     002    16-4W    CRX
DO YOU NEED HELP WITH THE SYMBOLS. PRESS<Y>IF YES
                                         <ENTER>OTHERWISE
  "Y"
THE CALL NUMBER IS 658P441P.  THE LIBRARY HAS 2 HOLDINGS
LINE 01:  COPY 1 IS IN STACKS--CIRCULATION DESK/STACKS 2ND
        FLOOR LIBRRARY.
THE LOAN PERIOD IS 16 WEEKS FOR FACULTY -  4 WEEKS FOR OTHERS
BUT IT IS CHARGED OUT AND RENEWED. AND IT IS DUE APRIL 27
        1986
LINE 02:  COPY 2 IS IN COMMERCE -- 101 LIBRARY
THE LOAN PERIOD IS 16 WEEKS FOR FACULTY - 4 WEEKS FOR OTHERS
```

Figure 3. Sample Title Search (part 2)

(i.e., B ST for browse subject topical). The results in this search do not match what the person wants. The interface then will try to search the general authority file (i.e., not just for topics but also geographical, persons, etc.). Failing to find anything, the interface automatically does a

```
PRESS <C> TO CHARGE OUT,<R> TO RENEW,<S> TO SAVE,<ENTER> TO
      GO ON

WISH TO SEARCH OTHER LIBRARIES? (PRESS<Y>IF YES,<ENTER>
     OTHERWISE

    "Y"

TYPE AN INSTITUTION NAME, CITY NAME, LIBRARY CODE, OR <ENTER>
    FOR ALL

TRITON COLLEGE

tls/peteprinc/tc

DEPAUL UNIVERSITY

tls/peteprinc/dp
```

Figure 3. Sample Title Search (part 3)

keyword title search (F T for find title). We took this approach because of the problems inherent in subject searching.[5] The results of the search in figure 4 were two records and the patron asked to see record one. The searcher makes another attempt to find headings by pressing the "H" key and is given the relevant heading "Ballistic Missile Defenses." The interface would then go back into the authority file under a new subject heading "Ballistic Missile Defenses." This later approach finds twenty-five additional books about "star war defenses." The assumption being made here is that the first subject heading is the most important. This then leads the patron to additional sources. When the user needs to link to an LCS record for circulation information, the interface takes care of this too.

```
TO BEGIN, TYPE A TERM AS GENERAL AS POSSIBLE TO DESCRIBE THE
    SUBJECT:
(YOU'LL BE ASKED TO PROVIDE MORE SPECIFIC INFORMATION IN A
    MOMENT. )
      "STAR WARS DEFENSE"
TYPE A MORE SPECIFIC WORD OR PHRASE AND <ENTER>.
( PRESS <ENTER> IF YOU AREN'T SURE OF THE WORDING. )
PRESS <ENTER> FOR ALL OR PRESS THE CORRESPONDING KEY IF THE
     SUBJECT
IS ABOUT A TOPIC--<T>, PERSON--<P>, CORPORATE--<C>, OR
     GEOGRAPHIC AREA--<G>
SEARCHING THE FULL BIBLIOGRAPHIC RECORDS OF THE HOLDINGS
     ACQUIRED SINCE 1975
b st star wars defense
                                             AUTHORITY  DISPLAY
      1.     STAR WAR FILMS--JUVENILE LITERATURE
      2.     STARA PAZOVA, SERBIA
      3.     STARAIA RUSSIA (R.S.F.S.R)--MUSEUMS.
      4.     STARBORN MAGIC MUSHROOM
      5.     STARCH
      6.          --MARKETING.
      7.          --PERIODICALS
```

Figure 4a. Example of a Subject Search (part 1)

The interface is menu-driven with an interactive dialogue between the user and the personal computer. An analysis of research on user aspects of computer designs proved to us the following:

1. a menu dialogue should be employed when the command set is so large that users are not likely to commit all commands to memory;

```
      8.    STARCH IN MEDICINE
      9.    STARCH INDUSTRY
      10.         --EQUIPMENT AND SUPPLIES

THESE ARE THE CLOSEST SUBJECT HEADINGS. DO ONE OF THE FOLLOWING
TYPE A NUMBER AND <ENTER> TO SEE THE CORRESPONDING
    BIBLIOGRAPHIC RECORDS
PRESS <B> TO BROWSE--TO SEE MORE HEADINGS
PRESS <E> TO END THIS SEARCH
PRESS <I> FOR AN INTERPRETATION OF THE SYMBOLS TO THE LEFT OF
    THE HEADINGS
IF THESE HEADINGS ARE NOT SATISFACTORY, PRESS <ENTER> TO TRY
ANOTHER SEARCH
t s star wars defense
t s star wars defense#
TRYING TO FIND STAR WARS DEFENSE IN TITLES...
f t star war defense
                                     BIBLIOGRAPHIC DISPLAY
1. BOVA, BEN, 1932- ASSURED SURVIVAL PUTTING THE STAR WARS
   DEFENSE IN PERSPECTIVE / BEN BOVA. BOSTON : HOUGHTON
   MIFFLIN, 1984. VII 343 P. ; OCM10-780039
```

Figure 4b. Example of a Subject Search (part 2)

2. a menu dialogue should be considered for inexperienced users because little training is needed;
3. a menu dialogue should be used when at least some of the users may be unfamiliar with the system functions; and
4. that the wording and order of any menu should be consistent with the command language.[6]

```
2. SHERR, ALAN B. LEGAL ISSUES OF THE STAR WARS DEFENSE PROGRAM
   / BY ALAN B. SHERR :BOSTON: : LAWYERS ALLIANCE FOR NUCLEAR
   ARMS CONTROL, :1984?:  38 P. ; OCM11-418172
FOUND 2 RECORDS
THESE ARE SHORT RECORDS 1 - 2.  DO ONE OF THE FOLLOWING:
PRESS <C> FOR CIRCULATION INFORMATION.
TYPE A NUMBER AND <ENTER> TO SEE THE CORRESPONDING FULL RECORD
PRESS <ENTER> TO END RECORD DISPLAY
"1"
s 1
    BOVA, BEN, 1932
ASSURED SURVIVAL PUTTING THE STAR WARS DEFENSE IN PERSPECTIVE /
  BEN BOVA. BOSTON : HOUGHTON MIFFLIN, 1984
   VIII  343 P.  ; 22 CM
   BIBLIOGRAPHY : P. 342-343.  ISBN 0395364051
1. BALLISTIC MISSILE DEFENSES--UNITED STATES  2. ATOMIC WARFARE
3. SPACE WEAPONS  4. UNITED STATES--MILITARY POLICY  I. TITLE

PRESS <H> TO MAKE ANOTHER ATTEMPT TO FIND RELEVANT HEADINGS
    OR PRESS <ENTER> TO GO ON

   "H"
  A RELEVANT HEADING IS: BALLISTIC MISSILE DEFENSES
```

Figure 4c. Example of a Subject Search (part 3)

We met these criteria by having large numbers of different commands, users who were inexperienced, and two systems with completely different functions. A microcomputer-based menu interface solved our problem by eliminating the need for commands. Also the two systems become transparent to patrons. The difficult thought processes presented in figure 1 were no longer a problem.

The Benefits of Microcomputers

Using microcomputers has also allowed us to gain more than just a user friendly system. The benefits are applicable to any other system even if it was vendor developed.

More Efficient Use of a Mainframe or Minicomputer

There has been a reduction in our error rates in FBR. Errors averaged almost 28 percent on the nonpersonal computer or dumb terminals and fell to only 6 percent on the intelligent terminals (see fig. 5). This meant that on the dumb terminals, one in every four searches resulted in an error message. LCS, which is not as complicated as FBR, has also seen a corresponding drop in the number of bad searches. Two other studies have found error rates of 11 and 13 percent respectively.[7]

	PERSONAL COMPUTERS	OTHER TERMINALS
LCS	1%	11%
FBR	6%	28%

ERROR RATES FROM OTHER STUDIES:

BERKELEY	11 %
OHIO STATE	13 %

Figure 5. Average Error Rates for 1985

In addition, the load on the mainframe is balanced. Some people will be sending searches while others are still formulating their search strategy on the personal computers. This guarantees that hundreds of simultaneous commands do not reach the front-end processor at the same time. Every search on a computer, no matter if it results in a good response or an error message, takes up machine resources. With 35 million searches in LCS, an error rate of 11 percent meant that over 3.5 million searches wasted computer resources. This level of erroneous searches cannot be an efficient use of any machine. If the mainframe or minicomputer is not large enough to handle the load, this error ratio can cause degradation of response time. Even with FBR and the introduction of more time-consuming keyword searches, we have not degraded response time.

Transparent Interface That Doesn't Require Learning System Commands

All searches, except subjects, go first to LCS because it contains records for everything cataloged at UIUC. Unsuccessful searches are then automatically routed to FBR. This routing takes place at the local level before the search goes to the mainframe. Commands are not input by the patron. Stopwords are not a problem. Explanations are given for all codes found in either LCS or FBR. Patrons are also led through a search of an authority file with cross and see references, then to the corresponding bibliographic record, and then to a separate database—LCS—which contains circulation information. The movement back and forth between FBR and LCS is invisible using the interface. The number of questions about how to use our system has dropped dramatically. Anyone using our library can search the online catalog for a known item or subject without needing help. Staff morale has improved tremendously.

Fast Interface and Short Interactions with Databases

Transmitting each line back and forth between a terminal and a distant mainframe can be slow. In addition to the communication distance, there is the possibility of slow response due to overloading the mainframe with additional searches. With a personal computer, nothing leaves the terminal until the search strategy is complete and correct. The initial communication is between the keyboard and the program in the personal computer and is therefore quite fast. Patrons know that the computer is working on their answer because the word "searching" blinks on the screen until a response appears.

Adaptable and Easily Changed Interface

It is advantageous to be able to improve the interface quickly as the system capabilities change. A local interface, using software that is purposely easy to update, helps accommodate system changes. It is also easy to test and fine tune a local interface. Our interface has gone through over thirty-five different versions in only three years. In addition, on a micro it is possible to have different types or levels of an interface on different terminals. There could be one version of the interface for undergraduates and visitors and another not as detailed for faculty offices.

The interface version available at UIUC allows the patron to charge out books. When a patron searches for a book, he is asked if he wants to charge out the book. The program asks for the patron's ID number and then automatically charges the book out. An explanation is given on what happens next and whom to contact if there is a problem. Other LCS sites do not allow patron charging of local books and therefore have a different version of the interface.

William Potter analyzed the effect our personal computers have had on borrowing materials from other LCS schools. Before the personal computers, interlibrary borrowing using LCS was 2.8 percent of UIUC's total circulation on LCS. After the introduction of personal computers, this figure jumps to over 8 percent. The increase in absolute numbers was from 35,182 items borrowed in 1982 by UIUC from other LCS schools to 123,123 in 1985.[8] This represents a 350 percent increase. By the interface asking patrons if they want to borrow from another library, resource sharing increased significantly. Until the personal computers were introduced, patrons did not realize that they had access to over 15 million volumes around the state. This same concept could apply when the Statewide ILLINET Catalog becomes operational.

Another creative use of our interface occurs at a local public library. A version of the interface, available on a personal computer at the Urbana Free Library, automatically dials into our system. These public library patrons get to search our system to locate what they cannot find locally. Since our system searches by keyword and subjects, they also have more access points than the CLSI terminal located nearby. This same concept is applicable to people with home computers and modems.

Increased Computing Power and Decreased Costs per Megabyte

When we purchased our microcomputers in 1983, they cost us approximately $1700 each for 128K and no disk drives. Today, a personal computer with two disk drives and a ten-megabyte hard disk costs around $1600. Although this cost is somewhat higher than a dumb terminal, the benefits in computing power, speed, and potential to access other systems far outweigh this difference. The era of the twenty- or thirty-megabyte hard drives is rapidly giving way to drives with gigabyte storage. The microcomputer of today is equal in power to and lower in price than many minicomputers of a few years ago.

We are at present investigating the feasibility of putting the statewide ILLINET Online Catalog on compact disc-read-only memory (CD-ROM) using the LePac system developed by Brodart.[9] Using compact disc for ILLINET could extend access to this valuable resource to even the smallest libraries. Compact disc technology would not require telecommunications hookups or charges, and expanding the network would not create the need for additional mainframe computer facilities. With the era of "write often, read many times" CD-ROM around the corner, many online systems could fit into a micro having the capability of storing 4 million MARC records.

Hardwired Microcomputers Can Search Multiple Databases and Catalogs

All of the eighteen regional system libraries in Illinois have access to our online catalog. They use the online catalog for interlibrary loan and bibliographic verification. Most of them also have their own online catalogs or circulation systems. Having to remember commands for their own systems and then our difficult command structure caused many problems.

Through an Illinois State Library grant, Cheng set up microcomputers in three different system libraries. These personal computers search our online catalog using his interface. Then, by pressing a function key, they instantly switch to their Data Phase, CLSI, or DRA local circulation systems to search using the command language. Two communications cards and changes in the interface make this switching between systems located on two different types of computers possible. They use the keyword and subject searching capability of our system and then search for the books on their systems. This occurs using the same microcomputer. With additional programming even this process could be automatic and searches could be saved and executed from system to system.

Microcomputers as Information Gatekeepers

A microcomputer with an 100-megabyte hard disk or gigabyte CD-ROM drive can store and search local or vendor-supplied databases. It is possible to purchase portions of databases from BRS, mount them on a mainframe, and search using a microcomputer. This same microcomputer could store the results of that search. This result could then be run against an online catalog to see if the local library owns the journals or reports.

We recently purchased InfoTrac for our Undergraduate Library. InfoTrac searches for magazine articles using an optical disc and a personal computer. At present, students must leave the InfoTrac terminal and proceed to a different terminal to search the online catalog. It would be far more beneficial if this same microcomputer could switch and search our online catalog. A patron would instantly know if we had the journal and its call number. We have already tied two different online systems together in the system libraries. This same microtechnology could also do this with InfoTrac and an online catalog. Having to search systems using different terminals would become obsolete.

An attempt at an end user searching system connected to the online catalog is being developed by our engineering librarian, Bill Mischo. Figure 6 is a prototype screen for a microcomputer-based system for librarians to automatically dial up and log on many systems.[10] Bill is adapting this prototype to allow students to automatically dial up some databases on BRS. After searching BRS, the results are run against our online catalog to determine if we own the items. All this will take place using an IBM-AT personal computer. This one terminal will be an information gatekeeper and will allow access to books and to various periodical indexes.

DATABASE VENDORS OR NETWORKS:

0. LCS/FBR
1. BRS—TELENET
2. BRS—TYMNET
3. DIALOG—TELENET
4. DIALOG—TYMNET
5. OCLC—TELENET
6. OCLC—TYMNET
7. RLIN—TYMNET
8. SDC—TELENET
9. SDC—TYMNET
10. RESUME SEARCHING—ALREADY ONLINE
11. PRINT PREVIOUSLY DOWNLOADED DATA
12. SIMULATION OF ONLINE SEARCH
13. LOGON BY SEARCHER
14. LCS VIA SWITCH
15. REFERENCE INFORMATION
16. KNOWLEDGE INDEX—TELENET
17. KNOWLEDGE INDEX—TYMNET

CHOOSE ONE OF THE NUMBERS

Figure 6. Illinois Search Aid for Expediting Online Database Searching

Although I painted a rosy picture of microcomputers and how they help make our system user friendly, there is one problem for dial-access patrons. Unless the dial-up user has an IBM personal computer, he or she will have to use commands to search our system. Since we support nine dial-up ports and two ports connected to a coaxial cable on campus, this could be a problem. However, the trade-off for having a user friendly system in the library *v.* the old system made the microcomputer our answer.

Conclusion

I began this discussion with a quote by McLuhan who said "If it works it's obsolete." That might be true of some technology but not completely true when discussing microcomputers. Any given machine, like the eight-bit machine, might become obsolete. However, this indicates that care be taken in choosing a machine that is state of the art. Another important consideration is that the microcomputer is expandable to take advantage of any future technological changes. A recent advertisement for the Compaq portable sums up the revolution taking place in microcomputers. It said: "Introducing the new Compaq Portable II—30% smaller, 17% lighter, and 400% faster!"[11]

Finagle's statement "The information you need is not available" is also rapidly becoming obsolete. It is possible that the information is available in one of over 3010 publicly available databases or within one of the 1.7 billion online records.[12] The problem that remains is how to make the public aware of this fact and how to allow them easy access.

Our experience with microcomputers has led me to the realization that they are the tool to unlock these bulging information storehouses. Just as our interface made resource sharing easy and increased interlibrary borrowing, so too could microcomputers act as information gatekeepers.

REFERENCES

1. Green, Jonathan, comp. *Morrow's International Dictionary of Contemporary Quotations*. New York: William Morrow and Co., Inc., 1982, p. 224.

2. Eriksen, Tore Linne. *The Political Economy of Namibia: An Annotated, Critical Bibliography*. Stockholm: The Scandinavian Institute of African Studies, 1985, p. i.

3. Bernard G. Sloan to ISL Automation Advisory Committee, personal communication, 7 April 1986.

4. Cheng, Chin-Chuan. "Microcomputer-Based User Interface." *Information Technology & Libraries* 4(Dec. 1985):346.

5. For an insight into this problem see: Cochrane, Pauline A., and Markey, Karen. "Catalog Use Studies—Since the Introduction of Online Interactive Catalogs: Impact on Design for Subject Access." *Library and Information Science Research* 5(Winter 1983):337-63; or Brownrigg, Edwin, et al. *Users Look at Online Catalogs: Results of a National Survey of Users and Non-Users of Online Public Access Catalogs. Final Report to the Council of Library Resources*. Berkeley: Division of Library Automation, University of California, 1982.

6. Williges, Beverly H. "User Considerations in Computer-Based Information Services" (Ada106-194-4). Springfield, Va.: NTIS, 1981, p. 31.

7. McPherson, Dorothy. "How the Melvyl Catalog is Used: A Statistical View." *DLA Bulletin* 5(Aug. 1985):18; and Norden, David J., and Lawrence, Gail H. "Public Terminal Use in an Online Catalog: Some Preliminary Results." *College & Research Libraries* 42(July 1981):313.

8. Potter, William Gray. "Creative Automation Boosts ILL Rates." *American Libraries* 17(April 1986):245.

9. For an explanation of LePac, see Schaub, John A. "CD-ROM for Public Access Catalogs." *Library High Tech* 3(Nov. 1985):7-11.

10. Mischo, William H. "Options for Subject Search Enhancement in Online Catalogs," 1985, p. 13 (typewritten).

11. Advertisement in *PC World* 3(May 1986):38-39.

12. Mischo, "Options for Subject Search Enhancement," p.17.

TAMAS E. DOSZKOCS

Computer Research Scientist
National Library of Medicine
Bethesda, Maryland

Natural Language User Interfaces in Information Retrieval

Introduction

This paper examines the role of natural language (NL) processing in information retrieval in the context of large operational information retrieval systems and services. State-of-the-art information retrieval systems combine the functional capabilities of the conventional inverted file—Boolean logic term adjacency approach—commonly employed by commercial search services, with statistical-combinatorial techniques pioneered in experimental information retrieval (IR) research, and formal natural language processing methods and tools borrowed from artificial intelligence (AI). The emergence and ever increasing importance of end-user searching provides challenging opportunities for the integration of sophisticated natural language analysis and processing techniques in user friendly interfaces.

IR systems achieve remarkable search speed and flexibility despite the virtual absence of formal language analysis procedures and meaning interpretation of the underlying text content.[1] State-of-the-art IR systems pragmatically blend the best features of diverse probabilistic-combinatorial and Boolean logic retrieval models and readily support free-form natural language user interfaces.[2] Yet, such direct natural language interfaces need to incorporate more sophisticated natural language processing and other Applied artificial intelligence techniques in order to cope intelligently with the inherent ambiguities of natural language queries and text, and to compensate for the inevitable semantic loss and confounding inherent in indexing and query matching.[3]

By the same token, AI approaches to natural language processing, particularly as applied to NL user interfaces and text searching, benefit from proven IR concepts and techniques in order to transcend still-prevalent domain-specificity and performance problems.[4] Thus natural language databases and user friendly online searching (UFOS) represent challenging and mutually supportive common problem areas for both IR and AI.

Information Retrieval

Although a variety of access methods have been developed for text retrieval,[5] operational IR systems are, almost without exception, characterized by the dominance of the inverted file, Boolean logic search paradigm.[6] In their basic form, IR systems are designed to manipulate fundamentally simple natural language text structures—such as bibliographic citations or full-text documents—although some IR software packages have been enhanced to incorporate generalized database management system (DBMS) access methods and special processing functions needed for the handling of integrated textual, numeric, graphics, and image data. The automatic "indexing rules" or algorithms are, by and large, lexically based procedures guided by delimiters and/or lists of nonindexed "stopwords," and the resulting inverted search keys are typically organized into B-tree structures for acceptable trade-off between speed of access and ease of updating. This approach is, for practical purposes, highly flexible and domain independent, although distinct indexing rules are needed in order to usefully fragment special textual fields, for instance, chemical names. In addition, in most IR systems, the automatically generated keywords are frequently augmented with human-assigned subject headings or thesaurus entries, mostly noun phrases, for added syntactic and semantic precision in searching.

Search queries, regardless of their form of input, must be ultimately transformed (typically, by trained intermediary searchers) to conform to the basic indexing scheme of the given IR system. Query search keys are matched against the inverted file search keys, and the corresponding ordered inverted lists are compared using exact-match Boolean AND, OR, AND NOT logic operators implemented as set operations (intersection, union, and complement, respectively) on the inverted lists. The process is exceedingly fast and efficient as implemented via currently available hardware and software architectures. The lack of linguistic and cognitive analysis procedures at indexing time and the resulting precision/recall problems are, to some extent, alleviated by the availability of powerful pattern-matching functions and metric (hierarchical/positional) operators, such as character masking, truncation, and adjacency. Trained intermediary searchers, in turn, provide the needed augmented "knowledge base," "inference engine," and "control strategy" for the proper configuration and sequencing of retrieval operations.

In the last few years, operational IR systems have been implemented that reflect the influence of G. Salton's seminal work and incorporate many useful ideas and results from theoretical and experimental IR research going back to the mid-1960s.[7]

IR research on the whole has been dominated by nonlinguistic, primarily lexical-statistical NLP methods as exemplified by Salton's SMART system[8] and the work of researchers such as Bookstein, Kraft, Rijsbergen,[9] and many others. Although experimental IR research prototypes have often suffered from problems of scale limitations in their technical approaches, IR research has nonetheless contributed a coherent conceptual framework and identified a core of desirable IR functions and evaluation methods that are of particular relevance to the design and implementation of end-user oriented information retrieval systems and services. The most important among these are the notions of (1) unrestricted natural language query input, (2) closest-match search strategy, (3) the ranking of the retrieval output according to expected relevance to the query, and (4) the dynamic utilization of user feedback in automatic query reformulation and search strategy modification.

Large-scale implementations of NL IR interfaces, such as CITE,[10] efficiently combine the best features of the Boolean and the probabilistic/combinatorial retrieval models with limited "intelligent" computational linguistic analysis and AI-type search heuristics. Such end-user oriented systems treat unrestricted natural language both in queries and text records as the least common denominator among different searchers, databases, and IR systems, offering the potential of true transportability and transparency among diverse users, information sources, and search systems. In order to successfully emulate the trained searcher's augmented retrieval "knowledge base," "inference engine," and "control strategy," however, UFOS must possess intelligent NLP capabilities of greater sophistication.[11]

The convergence of a number of hardware, software, and user trends necessitate the augmentation of conventional information retrieval and filtering capabilities with appropriate and efficiently implemented techniques adapted from AI application areas—such as natural language processing and understanding, expert systems, intelligent information management, and intelligent problem-solving. The trends include: full-text databases; very large databases; mixed information sources containing text, numerics, graphics, image, and other data; the increased diversity, depth, and breadth of information sources; hybrid technologies (e.g., compact disc-read-only memory, CD-ROM); special-purpose IR hardware; non-von-Neumann computer architectures; associative memories; distributed and parallel processing; and interactive end-user searching and personal, computer-enhanced "information metabolism."

Artificial Intelligence

Artificial intelligence, particularly applied artificial intelligence, is a growth industry of high expectations.[12] Notwithstanding the difficulty of defining intelligence, let alone artificial intelligence, and despite the many lingering doubts and the seeming tower of Babel situation, the AI field possesses a healthy basic research underpinning and the AI community has been successful in developing important concepts, tools, techniques, and applications of interest to other information-intensive disciplines.[13] Just as "conventional" programming languages and diverse data representation models in "classical" IR and DBMS systems serve as problem-solving tools for a wide variety of applications, AI knowledge representation models, such as production rules,[14] frames,[15] and semantic networks as well as AI programming languages, such as LISP (list processing), PROLOG (logic programming),[16] OPS5 (rule-based programming),[17] SMALLTALK (object-oriented programming), and other general purpose techniques and tools (e.g., heuristic search, ATN [automatic translation network] parsers, and rapid prototyping,[18]) represent versatile AI implements for a broad range of applications including IR.

Natural Language Processing—Problems

NLP and computational approaches to dealing with natural language were among the earliest objects of interest in AI research. The well-known ambitions and subsequent failures of machine translation research and development projects in the 1960s, as well as the recent success of more limited yet highly pragmatic computer-assisted language translation systems, epitomize both the great difficulties involved in coping with natural language and the degree of maturation and realistic goal setting of the field.

Understanding language, though it appears to humans to be naturally easy, is a difficult task that involves highly cognitive and not yet totally understood intellectual and psychological processes.[19] Schank argues that NLP is both a linguistic and a cognitive task, and it cannot occur without a knowledge base concerning the relevant subject area.[20] This in turn points to a synthesis of natural language processing and expert systems techniques. Typical areas of concern and investigation in AI NLP research involve automated lexical, syntactic, and semantic analysis; dealing with ill-formed or fragmentary input; ellipsis; conjunction and negation; pronoun/anaphora resolution; definite noun phrases, quantification; beliefs and intentions; fail-soft recovery; space and time and contextual understanding—to name just a few.

The role of NLP in information retrieval is less clear[21] given that the very nature of the textual documents traditionally dealt with in IR does not warrant elaborate analysis, and the fundamentally mechanistic techniques developed in IR have served to handle the passive query-to-document matching problem at an acceptable level of success. The advent of unmediated "user friendly" search interfaces on the one hand, and the considerably broader scope and depth of today's full-text databases on the other, however, necessitate a careful reexamination of this role.

Computational linguistic processing has reached fairly reliable stability and practical utility in morphological and syntactic analysis. While procedures for morphological analysis are decidedly nontrivial, they are generally more straightforward and efficient than syntactic analysis. Common strategies for morphological analysis employ some or all of the following: morphological rewriting rules, dictionary lookup, inflection generator, complexity testing, and idiom and compound recognition.[22] Prefixing and suffixing are the main concern in processing English. Often, in processing specialty languages—e.g., medical English—special attention must be paid to characteristic prefixing, suffixing, and morphosemantic problems.[23] Compare, for instance:

ALEXIA—DYSLEXIA
VITAMINS—AVITAMINOSIS
ANXIETY—ANTIANXIETY AGENTS
INFECT—DISINFECT
HYPERTENSION—HYPOTENSION
ADJUSTMENT—MALADJUSTMENT
ANTIBIOTICS—"-MYCINS"
INFLAMMATION—"-ITIS"

Reasonably—i.e. relatively—robust and efficient NL syntactic parsers have been developed and incorporated into many artificial intelligence applications. "Parsing efficiency is crucial when building practical natural language systems. This is especially the case for interactive systems, such as natural language database access, interfaces to expert systems and interactive machine translation."[24] Tomita's LR context-free parsing algorithm, for instance, takes advantage of the left-to-rightness of natural-language user input, and parsing starts as soon as the user types the first word, thus reducing apparent response time from the user's point of view.

Despite significant advances, a number of unsolved problems and limitations in AI NLP remain. Most importantly, none of the systems developed to date are fluent in the use of unrestricted natural language.[25]

For potential IR use of AI NLP techniques, it is important to remember that while most full sentences are unambiguous, their component parts are frequently ambiguous. Since IR systems utilize component or fragmentary lexical, syntactic, and semantic units as a result of indexing, and since indexing inevitably implies omission, the disambiguation problem in domain-independent IR systems is much less tractable than in narrow-domain artificial intelligence applications.

Acoustic-phonetic, lexical, syntactic, and pragmatic language complexity, as reflected in word-sense ambiguity, structural ambiguity, and the referential ambiguity of noun phrases, pose very challenging problems in user friendly artificial intelligence and information retrieval interfaces. Few if any existing NLP systems, for instance, are able to recognize and disambiguate compounds (compare database *v.* data base), acronyms (compare G-SUIT), or abbreviations (compare MD ==> Maryland ==> physician) in large textual databases.

Noun-phrase ambiguity may be further compounded in keyword indexing by ignoring word order, field, and other text boundaries, and by partial match search strategy in query matching, e.g.,:

MEDICAL LITERATURE ==> MEDICINE IN LITERATURE ==> LITERATURE IN MEDICINE

SEXUAL PERVERSION ==> SEXUAL ABSTINENCE

There is a lot of humor in all of this, as in:

HUMOR IN MEDICINE ==> AQUAEOUS HUMOR IN MEDICINE

or, put differently, language understanding can be a joke. For proper perspective it is worth keeping in mind that there still does not exist any NLP computer program that can handle nearly all of English syntax. None can even come close to coping with the semantics of all of the English language and none is on the horizon. "Wretched and confusing prose can defeat even human comprehension."[27]

The impossibly large number of rules that would be necessary for the morphological, syntactic, and semantic disambiguation of natural language in real-life multidisciplinary knowledge domains precludes using the rule-based expert system approach as well. (No accurate estimation has been made of how many context-free rules are needed to cover English almost completely, but the number is very large.)[28] From the survey of the literature it appears that artificial intelligence is not quite ready to field any system flexible enough for mass use, save a few relatively small problem domains.[29]

Can information retrieval facilitate finding artificial intelligence solutions to practical natural language processing problems? Certainly so!

Domain-independent IR search techniques can serve as efficient filters for more refined in-depth natural language processing. As more and more powerful AI concepts and tools become available (compare AI PC toolkits) to more and more end users, the problems of scale and performance limitations that characterize contemporary AI systems will be gradually overcome. To the extent that the AI and IR communities will be able to learn from each other and gain mutual insights into the problems of language and searching, there will exist intelligent IR systems and effective domain-independent AI NL search systems. And perhaps—in theory at least—the distinctions between the two (mind) sets will be "fuzzy" at best.[30]

Natural Language and User Friendly Online Searching

Natural language interface technology represents a major breakthrough in "user friendly" computer systems.[31] Along with other end-user-oriented interface techniques (such as menus, windowing, graphics, icons, pointing, touching) commercial implementations of NL interfaces (such as Artificial Intelligence Corporation's INTELLECT or Texas Instrument's NLMenu DBMS front-end products) target the largest hitherto untapped segment of the information marketplace, namely end users. The same holds true for public access online catalogs in libraries and for user friendly IR interfaces in general,[32] and for natural language information retrieval interfaces in particular (e.g., the National Library of Medicine's CITE system for searching the world's largest medical literature databases, MEDLINE and CATLINE).[33]

Online systems and the personal computer revolution have made computer resources universally available. Now a similar revolution in software (i.e., user interfaces) is needed to make the computer universally usable.[34] Natural language interfaces in information retrieval and artificial intelligence are the scouts, shock troops, vanguards, and sometimes martyrs of the user interface revolution.

With occasional end users already outnumbering trained professional searchers in the user populations of online information utilities like The SOURCE and CompuServe, it becomes increasingly important to develop and refine analytical cognitive models to better assess the user's skills and understanding of the information stored, and to match the user's cognitive model to the system's model and knowledge representation.

Of course natural language is not always natural for a user interface,[35] but it is particularly well suited for IR and DBMS interfaces due to the very large number of potential users, high volume of query transactions, distribution of costs over large numbers of users, and the fundamentally linguistic nature of the user-system information exchange.[36]

Ease of learning, ease of use, and transportability are among the most attractive features of natural language front-ends. Natural language shifts the burden of understanding from the user to the system thus allowing the user to focus on the problem at hand.[37] At the current state-of-the-art, the high development cost and less than 100 percent reliability of knowledge-based domain-specific NL DBMS systems (compare EXPLORER, developed by Cognitive Systems, Inc. for oil exploration) appear to have palled their widespread development and commercial use. Considerably more commercial success has been achieved by domain-independent NL DBMS front-ends. INTELLECT, for example, substitutes knowledge of the physical and logical structure of the database and the topology of user interactions and system functions for domain-specific semantic knowledge, using the inverted DBMS index as lexical pointers to the system's rather small domain-specific and semantically augmented dictionary. Despite its shortcomings in ambiguity resolution,[38] INTELLECT succeeds in its practical utility, reliability, and operational performance.

It is interesting to note that while a DBMS NL front-end like INTELLECT has to cope with the full spectrum of linguistic processing problems of a "habitable subset" of the English language—namely the limited domain of the DBMS command language and a relatively small number of query interaction paradigms—it does not have to deal with language ambiguity and matching problems at the level of the database content, due to the fact that DBMS systems deal, for the most part, with discrete and finite nontextual data.

By contrast, a natural language information retrieval textual database interface, like NLM's CITE system[39] must focus on language ambiguity and matching problems for all of English. At the same time, CITE need not invest a great deal of effort in full-scale linguistic analysis in query-to-command language translation due to the relatively small number of IR commands available and the simplicity of the underlying database structure. (It would be perfectly feasible and appropriate to use INTELLECT-like NLP techniques, instead of menu choices, in CITE to "understand" the type of query at hand, identify its topical component, as well as any implied or explicitly stated limitations as to type of material desired, language restrictions, currency of material.)

Natural Language Queries in NL Databases

Natural language information retrieval interfaces must then deal with the problem of language ambiguity at the level of both the query and the database content and must resolve the matching problem between free-form queries and the database in a manner acceptable to end users who typically approach a natural language system with high expectations of (artificial machine) intelligence. The basic problem areas of matching can be conveniently divided into lexical, syntactic, semantic, and special concerns.

Lexical Problems

The inverted keyword index of a NL database of 1 million records is likely to contain in excess of a quarter million distinct lexical words.[40] The frequency distribution of these lexical entries will typically conform to the characteristics of the Zipf distribution. This empirical fact has the following practical implications:

1. There will be a few dozen to a few hundred highly posted (i.e., high database frequency) index entries, many of which will be natural candidates for exclusion from the index (compare "stopwords").
2. Approximately half of all the index entries will have a database frequency of one, and as many as half or more of these may be misspellings. This suggests the incorporation of efficient spelling error detection/correction algorithms and dictionaries in the NL user interface and in indexing, with the dictionary preferably derived from and dynamically updated in conjunction with the database itself.
3. Knowledge of the frequency distribution of the lexical entries suggests implicit heuristics for automatic search strategy formulation. The NL interface must be capable of recognizing and properly dealing with frequent acronyms, abbreviations, numerals, chemical names, names of people/syndromes:

X RAY ==> X-RAY; CAT SCAN; US ==> USA ==> U.S. ==> U.S.A.; AI ==> ARTIFICIAL INTELLIGENCE; VITAMIN B 1 ==> VITAMIN B1 ==> THIAMIN; TYPE 1 ==> TYPE I; FACTOR V; 2,4,5-T ==> TRICHLOROPHENOXYACETIC ACID ==> AGENT ORANGE; EPSTEIN-BARR VIRUS; BARR BODIES; BARRE-LIEOU SYNDROME; GUILLAN-BARRE SYNDROME

orthographic and phonetic transcribing and transliteration:

TUMOUR ==> TUMOR;
GYNAECOLOGY ==> GYNEKOLOGIA ==> GYNECOLOGY

idioms and clichés:

OFF COLOR; CHANGE OF HEART; OUT OF SIGHT

slang, lingo, jargon, and lore:

POT; SPEED; ANGEL DUST; CRACK

stopwords that are noun-phrase components:

VITAMIN A; HEPATITIS A; "TO BE OR NOT TO BE";
THE "ME" GENERATION

compound words:

DATABASE ==> DATA BASE; ONLINE ==> ON-LINE ==>
ON LINE; BACKACHE ==> BACK ACHE ==> BACK PAIN

Morphological Analysis

Morphological analysis (stemming or "conflation") warrants special attention in large-scale natural language information retrieval interfaces. The automatic identification of lexical roots in inverted file based operational IR systems is the first step in the process of matching query words and inverted index entries. The latter may have been derived from auxiliary vocabularies that serve as semantic database navigational tools, or from the database records themselves. The following examples from the MEDLINE and MEDICAL SUBJECT HEADINGS inverted indexes illustrate ever-present and familiar lexical ambiguities and semantic noise introduced as a result of using word roots with variable-length masking operations in matching:

ACCESS ==> ACCESSORY
ASPIRIN ==> ASPIRATION
AUDIT ==> AUDITORY
BATTERED ==> BATTERY
COMMUNICABLE ==> COMMUNICATIONS
CREATINE ==> CREATIVENESS
DIGITAL ==> DIGITALIS
EXTREME ==> EXTREMITIES
EXPECTATION ==> EXPECTORANT
INFANT ==> INFANTILISM
INFORM ==> INFORMAL ==> INFORMER
LABOR ==> LABORATORY
MEDIA ==> MEDIAN ==> MEDIATED
METHOD ==> METHODIST
MIGRAINE ==> MIGRANT
NURSERY ==> NURSES ==> NURSING
RECEPTION ==> RECEPTORS
SHORT ==> SHORTAGE ==> SHORTHAND
TREAT ==> TREATMENT ==> TREATISE ==> TREATY

Short words require even more caution:

AID ==> AIDS
ANAL ==> ANALYSIS
APE ==> APES ==> APEX
ARM ==> ARMY
CARE ==> CAREER
FAIR ==> FAIRS
HEAR ==> HEARING ==> HEART ==> HEARTWATER ==> HEARTWORM

Many thousands of other such examples could be found.

Syntax-Related Problems

Natural language topical surrogates—e.g. book and journal titles, headlines, table of contents, back-of-the-book indexes—are usually expressed via larger syntactic units, mostly noun phrases. Noun phrases are frequently ambiguous in and of themselves—e.g., SHELLFISH POISONING ==> POISONING OF SHELLFISH ==> POISONING BY SHELLFISH. The use of Boolean operators, metric operators, and *m* out of *n* weighted-logic closest-match search strategy—in order to compensate for the lack of linguistic analysis in indexing—further compounds the text-matching problem. Consider, for example:

ABUSE ==> CHILD ABUSE ==> DRUG ABUSE ==> ELDER ABUSE ==> SPOUSE ABUSE

CRISIS MANAGEMENT ==> MANAGEMENT CRISIS ==> MANAGEMENT BY CRISIS ==> CRISIS BY MANAGEMENT

KIDNEY ==> KIDNEY BEAN LECTINS ==> KIDNEY DISEASES

SEXUAL ABSTINENCE ==> SEXUAL PERVERSION

SHORT TERM EFFECTS ==> SHORT TERM MEMORY ==> SHORT TERM PSYCHOTHERAPY

Since literally hundreds of similar lexical and syntactic matching problems are encountered daily in a large operational NL IR system, it is evident that automatic query analysis and matching can substantially benefit from morphological and syntactic analysis in order to lend additional precision to the available truncation, character masking, Boolean, metric, weighted-logic, or generalized pattern-matching strategies. Consider for instance the automatic generation of Boolean search statements from NL queries.[41]

Semantic Problems

A great many formal semantic aids and ad hoc heuristics are used by trained searchers when interacting with information retrieval systems. Some examples are controlled vocabularies, "hedges," "preexplodes," multidatabase cross indexes, stored search strategies, and the like. Systems that rely on controlled vocabularies often lack in currency, database warrant, or conceptual exhaustivity. For instance, *Medical Subject Headings* does not currently (1986 edition) have a subject heading for BIOTECHNOLOGY, and it uses PMS as a cross reference to an older subject heading PREGNANT MARE SERUM (GONADOTROPINS, EQUINE), but at the same time PMS is not linked to PREMENSTRUAL SYNDROME, nor is AI linked to ARTIFICIAL INTELLIGENCE.

As was noted earlier, systems without automated semantic aids shift the full burden of query understanding and matching on the searcher. NL IR interfaces must minimally rely on and must intelligently utilize existing machine-readable semantic search aids. The existing aids, however, need to be augmented by additional semantic mapping tools such as statistical term associations, switching vocabularies, enriched "fuzzy" thesauri, "scriptal" micro-lexicons, production rules, and/or heuristics elicited from expert searchers.

The natural language information retrieval interface must also be designed to deal with special problems such as multiple languages, specialty languages, dialects, and professional jargon.[42] To a considerable extent, the CITE experimental R & D system and its operational versions have attempted to address many of the linguistic problem areas outlined in this paper.[43]

CITE represents a domain-independent NL IR interface approach that combines conventional inverted file, Boolean logic with term frequency-based weighted logic, closest-match search strategy and efficient NLP techniques involving "intelligent" stemming,[44] partial syntax analysis, automatic query-to-controlled-vocabulary mapping, look ahead search ambiguity resolution and filtering of combinatorial controlled vocabulary term displays, automatic user feedback processing, and other techniques adapted from applied AI such as domain-specific semantic navigational tools, refined textual pattern matching, and ad hoc expert searcher heuristics. The appendix illustrates several NL user interactions on the CITE system. The last example serves to put things in humbling perspective: The search query NATURAL LANGUAGE PROCESSING automatically picks up the subject heading NATURAL DISASTERS!

Other AI Applications of Direct Relevance to IR

In addition to natural language processing techniques in general and NL DBMS front-ends in particular, the following artificial intelligence areas are perceived by this author to be of direct relevance to IR:

1. *Expert systems.* To the extent that rule-based expert systems are instances of very high level programming tools that allow the expression of order-independent rules instead of ad hoc pieces of order-dependent conventional program code, they can be of benefit in NL IR interface development in capturing trained searcher expertise as well as codifying broad linguistic processing rules. Efficient microcomputer implementations of rule-based expert systems are becoming increasingly available.[45]
2. *Intelligent information management.* Intelligent information management involves the analysis of the interrelationships among multiple databases, information sources, user behavior, including observation of past actions and codified procedures, in order to develop rules for enhanced data retrieval and management.[46] Elements of this approach have been utilized in information retrieval—e.g., in the PAPERCHASE system[47] and the Syracuse University SUPARS Project. The systematic development and utilization of intelligent information management techniques should benefit IR systems in the future.
3. *AI knowledge representation techniques.* In general, AI researchers have found that amassing large amounts of knowledge rather than sophisticated reasoning techniques are responsible for the power of expert systems.[48] The knowledge encoded in conventional controlled vocabularies can be potentially augmented via rule-based relationships, "ISA" knowledge-representation constructs,[49] or predicate calculus statements and fuzzy logic.
4. *Integrating IR, DBMS, AI, and other technologies.* To date, relatively little work has been done in such integrative R & D. Videotex[50] and CD-ROM systems are perhaps the most promising applications in this category. The latest videotex systems—e.g. The SOURCE and CompuServe—combine sophisticated IR, DBMS, electronic mail, and other technologies, and typically offer full-text inversion and Boolean search, distributed database management as well as menu-driven, tree-structured, user friendly access. CD-ROM database publishing and special-purpose knowledge-base applications similarly combine state-of-the-art IR, DBMS, AI, and overall information-management technologies. The integration of efficient NLP techniques and intelligent computer-assisted instruction[51] capabilities will enable videotex and CD-ROM users, and users of diverse information utilities to augment their own intellectual power by machine intelligence that is perhaps

going to be able to grasp—if not understand—database information content, discover new relationships, synthesize new knowledge, and postulate new hypotheses.

REFERENCES

1. Salton, Gerard, and McGill, Michael. *Introduction to Modern Information Retrieval.* New York: McGraw-Hill, 1983; and Sparck-Jones, K., and Kay, M. *Linguistics and Information Science.* New York: Academic Press, 1972.

2. Doszkocs, Tamas E., and Rapp, Barbara A. "Searching MEDLINE in English: A Prototype User Interface with Natural Language Query, Ranked Output, and Relevance Feedback." In *Information Choices and Policies* (Proceedings of the ASIS 42nd Annual Meeting, Minneapolis, Minn., 14-15 Oct. 1979), edited by Roy D. Tally and Ronald R. Deultgen, pp. 131-39. White Plains, N.Y.: Knowledge Industry Publications, 1979; Koll, M., et al. "Enhanced Retrieval Techniques on a Microcomputer." In *The National Online Meeting* (Proceedings of the 5th National Online Meeting, New York, 10-12 April 1984), compiled by Martha E. Williams and Thomas H. Hogan. Medford, N.J.: Learned Information, 1984; and Berstein, L.M., and Williamson, R.E. "Testing of a National Language Retrieval System for a Full Text Knowledge Base." *JASIS* 35(July 1984):235-47.

3. Doszkocs, Tamas E. "Natural Language Processing in Intelligent Information Retrieval." In *Proceedings of the ACM Annual Meeting*, edited by S. Ron Oliver, pp. 356-59. New York: Association for Computing Machinery, 1985.

4. Andriole, Stephen, J., ed. *Applications in Artificial Intelligence.* Princeton, N.J.: Petrocelli Books, 1985.

5. Faloutsos, Christos. "Access Methods for Text." *Computing Surveys* 17(March 1985):49-74.

6. Salton, and McGill, *Introduction to Modern Information Retrieval.*

7. Doszkocs, "Searching MEDLINE in English," pp. 131-39; Koll, et al., "Entrance Retrieval Techniques," pp. 165-70; and Bernstein, and Williamson, "Testing of a Natural Language Retrieval System," pp. 235-47.

8. Salton, Gerard, ed. *The SMART Retrieval System: Experiments in Automatic Document Processing.* Englewood Cliffs, N.J.: Prentice-Hall, 1971.

9. Bookstein, A. "Implications of Boolean Structure for Probabilistic Retrieval." In *Proceedings of the 8th Annual International ACM SIGIR Conference* (Montreal, Canada, 5-7 June 1985), edited by S. Ron Oliver, pp. 11-17. New York: Association for Computing Machinery, 1985.

10. Doszkocs, Tamas E. "CITE NLM: Natural-Language Searching in an Online Catalog." *Information Technology and Libraries* 2(Dec. 1983):364-80.

11. ________ , "Natural Language Processing," pp. 356-59.

12. Andriole, *Applications in Artificial Intelligence.*

13. Lunin, L., and Smith, Linda, eds. "Perspectives on Artificial Intelligence: Concepts, Techniques, Applications, Promise." *JASIS* 35(Sept. 1984):277-319; Grishman, R. "Natural Language Processing." *JASIS* 35(Sept. 1984):291-96; Cooper, W.S. "Bridging the Gap Between AI and IR." In *Research and Development in Information Retrieval* (Proceedings of the 3d Joint BCS and ACM Symposium), edited by C.J. van Rijsbergen, pp. 259-65. Cambridge: Cambridge University Press, 1984; Sparck-Jones, K. "Natural Language Access to Databases: Some Questions and a Specific Approach." *Journal of Information Science* 4(March 1982):41-48; and Kolodner, J. "Indexing and Retrieval Strategies for Natural Language Retrieval." *ACM Transactions of Database Systems* 8(1983):434-64.

14. Hayes-Roth, Frederick. "Rule-Based Systems." *Communications of the ACM* 28(Sept. 1985):921-32.

15. Fikes, Richard, and Kehler, Thomas. "The Role of Frame-Based Representation in Reasoning." *Communications of the ACM* 28(Sept. 1985):904-20.

16. Politt, A.S. "A 'front-end' System: An Expert System as an Online Search Intermediary." *Aslib Proceedings* 36(May 1984):229-34.

17. Brownston, Lee, et al. *Programming Expert Systems in OPS5*. Reading, Mass.: Addison-Wesley, 1985.

18. Schutzer, Daniel. "Artificial Intelligence-Based Very Large Data Base Organization and Management." In *Applications in Artificial Intelligence*, pp. 251-78.

19. Hendrix, Gary G., and Sacerdoti, Earl D. "Natural Language Processing: The Field in Perspective." In *Applications in Artificial Intelligence*, pp. 149-92.

20. Schank, Roger, and Schwartz, Steven P. "The Role of Knowledge Engineering in Natural Language Systems." In *Applications in Artificial Intelligence*, pp. 193-212.

21. Sparck-Jones, and Kay, *Linguistics and Information Science*.

22. Kay, Martin. "Morphological and Syntactic Analysis." In *Linguistic Structures Processing*, edited by A. Zampolli, pp. 131-234. New York: North-Holland, 1977.

23. Pacak, M.G., and Dunham, G.S. "Computers and Medical Language." *Medical Informatics* 4(1979):13-27.

24. Tomita, Masaru. *Efficient Parsing for Natural Language*. Hingham, Mass.: Kluwer Academic Publisher, 1986, p. xvii.

25. Hendrix, and Sacerdoti, "Natural Language Processing," pp. 149-92.

26. Golden, F.L. *Jest What the Doctor Ordered. A Recording of Medical Humor*. New York: Frederick Fell Publishers, 1949; and Cowan, L. and Cowan, M. *The Wit of Medicine*. London: Frewin, 1972.

27. Charniak, Eugene, and McDermott, Drew. *Introduction to Artificial Intelligence*. Reading, Mass.: Addison-Wesley, 1985.

28. Tomita, *Efficient Parsing*.

29. Hice, Gerald F., and Andriole, Stephen J. "Artificially Intelligent Videotex." In *Applications in Artificial Intelligence*, pp. 295-312.

30. Schmucker, Kurt J. *Fuzzy Sets, Natural Language Computations, and Risk Analysis*. Rockville, Md.: Computer Science Press, 1984.

31. Schank, and Schwartz, "The Role of Knowledge Engineering," pp. 193-212.

32. Cowan, and Cowan, *The Wit of Medicine*.

33. Doszkocs, Tamas E. "From Research to Application: The CITE Natural Language Information Retrieval System." In *Proceedings of the Fifth BCS and ACA SIGIR Conference*. Berlin, Germany: Springer-Verlag, 1983, pp. 251-62.

34. Carbonell, Jaime G. "The Role of User Modeling in Natural Language." In *Applications in Artificial Intelligence*, pp. 213-26.

35. Rich, E. "Natural Language Interfaces." *Computer* (Sept. 1984):39-47.

36. Petrick, S.R. "On Natural Language Based Computer Systems." In *Linguistic Structures Processing*, pp. 313-40; and Schank, and Schwartz, "The Role of Knowledge Engineering," pp. 192-212.

37. Carbonell, "The Role of User Modeling," pp. 213-26.

38. Schank, and Schwartz, "The Role of Knowledge Engineering," pp. 193-212.

39. Doszkocs, "CITE NLM," pp. 364-80.

40. Doszkocs, Tamas E. "AID—An Associative Interactive Dictionary for Online Searching." *Online Review* 2(June 1978):163-73; and Doszkocs, Tamas E., et al. "Analysis of Term Distribution in the TOXLINE Inverted File." *Journal of Chemical Information and Computer Sciences* 16(1976):131-35.

41. Salton, Gerard, et al. "Automatic Query Formulations in Information Retrieval." *JASIS* 34(July 1983):262-80.

42. Cowan, and Cowan, *The Wit of Medicine*; Bove, A., et al. "Hellenic Influence in Medical English." *Med Clin* 83(1984):209-13; and Burnum, J.F. "Dialect is Diagnostic." *Annals of Internal Medicine* 100(June 1984):899-901.

43. Doszkocs, "Natural Language Processing," pp. 356-59.

44. Ulmschneider, John E., and Doszkocs, Tamas E. "A Practical Stemming Algorithm for Online Search Assistance." *Online Review* 7(1983):301-18.

45. Lehner, and Barth, "Expert Systems on Microcomputers," pp. 109-24.

46. Schutzer, "Artificial Intelligence-Based Very Large Data Based Organization and Management," pp. 251-78.

47. Horowitz, G.L., and Bleich, H.L. "Paperchase: A Computer Program to Search the Medical Literature." *New England Journal of Medicine* 305(15 Oct. 1981):924-30.

48. Gevartner, William B. "Expert Systems: Limited but Powerful." In *Applications in Artificial Intelligence*, pp. 125-42.

49. Rada, Roy, et al. "A Medical Informatics Thesaurus." In *Proceedings of the MEDINFO '86 Conference* (26-30 October 1986, Washington, D.C.), pp. 1164-1172. North Holland: Amsterdam, Holland.

50. Hice, and Andriole, "Artificially Intelligent Videotex," pp. 295-312.

51. Fletcher, J. Dexter. "Intelligent Instructional Systems in Training." In *Applications in Artificial Intelligence*, pp. 427-52.

DAVID E. TOLIVER
Manager, Software Development
Institute for Scientific Information

Design Issues in Automatic Translation for Online Information Retrieval Systems

Introduction

One objective of computer intermediary systems is to minimize incidental and accidental differences among the many distinct languages found in online bibliographic retrieval. Three classes of languages are identified: access protocols, retrieval commands/responses, and database structures. Each class has its own characteristic requirements for automatic translation. In developing one intermediary product—the Sci-Mate Searcher—distinct translation approaches proved most effective for each class: a procedural language for access protocols, customized coding for retrieval commands/responses, and a knowledge-based table for database structures. Despite differences in translation methods, users are presented with a consistent view throughout the product.

The Problem: Online Babel

Online bibliographic information retrieval, from a systems point of view, is not user friendly. Using many heterogeneous online bibliographic services can be difficult for professional searchers and nearly impossible for occasional end users for several reasons. These include the problems of database selection, strategy development, and the overwhelming and sometimes contradictory details of usage and syntax. This paper addresses this last source of difficulty, one that is more likely to be solved in the near future by automation than the semantic and subject-knowledge issues.

Online services provide access to enormous amounts of information but at the same time pose linguistic barriers to their own broad usage. The number of services with distinct protocols and languages continues to grow. There are now five major packet switching networks in the United

States and one in every European country. At least fifteen bibliographic database hosts in the United States, Canada, and Europe offer hundreds of databases with many of the specialized databases found only on a single host. Besides containing unique information, each database is structured with distinct field designations and data coordination conventions.

The linguistic conventions used in online searching are by design terse and cryptic. With cost a function of time spent online, brevity is mandatory. For searching to be cost effective at even the relatively fast transfer rate of 2400 baud, a minimal user environment is preferable. However, this does not excuse the great diversity and incompatibility of commands, codes, and conventions.

From one system to the next, a given function usually is invoked by a keyword entirely unique to the system. With a few exceptions, there is no opportunity to define synonyms or otherwise improve consistency among the systems used. The babel of distinct protocols and language conventions now being used by online systems derives from the history of their development.

The packet switching networks—Telenet, Tymnet, Uninet, and Infonet—were developed independently from one another as competitive services. Each interacts with users using their own distinctive protocols and conventions.

In the early and mid-1970s, development efforts in bibliographic search software were independently conducted by several firms, notably Dialog (Lockheed), Orbit (SDC), and Bibliographic Retrieval Service (BRS). Dialog evolved out of Recon, funded by the National Aeronautics and Space Administration, while Orbit evolved out of Elhill under contract with the National Library of Medicine. BRS Search, originally derived from IBM's STAIRS software, has always been a commercial search service. During this independent development, there was no coordination of language terminology and syntax.

Many of the early commercial bibliographic databases derived from federal and private publishers of printed tertiary indexes. It is remarkable that data from so many diverse sources were brought together and made to work under one or more vendors' retrieval software. The data in many cases were not initially intended for distribution as an online database. It is quite understandable that most databases follow their own distinctive indexing and fielding conventions.

At the present time, economic and technical constraints work against significant change in the online systems' software. The networks, retrieval hosts, and database publishers all have invested thousands of man-years in software and data. Customers who have learned to use these systems depend on them remaining stable. All involved are understandably reluctant to look for and convert to any proposed standard, such as the European Common Command Language.[1]

Given this situation, the information retrieval specialists often must make choices that are not fully satisfactory. The specialists may limit themselves to one or two host systems and a few select databases. The capabilities of the systems and the databases' content and structure can thus be thoroughly mastered to accomplish all that is possible with the selected facilities. But in so restricting themselves, vital information on hosts and databases not used will be missed.

On the other hand, specialists may choose to learn how to use a broad selection of host systems and databases in order to access all possible sources of relevant information. Much of what must be learned are details of access protocols, commands, and database syntax—the only means to getting the information itself. Learning and maintaining proficiency with many different systems is costly and may result in specialists with more diffuse and less expert skills than those who choose a more limited scope.

Technical end users are even more restricted by the linguistic barriers posed by conventional database systems. They can afford less time than specialists to devote to learning details of access, retrieval, and database structure. They are frequently bewildered by the diversity of options. Most of the time, end users will turn to specialists to meet their information needs even though they know the technical terminology better and are better equipped to judge the results of the search.[2]

A Natural Solution: Intermediate Translating Computers

For more than ten years, various automated solutions to the linguistic problem have been proposed and implemented. These usually consist of a computer placed between users and one or more host systems. Known generally as "computer intermediaries," these systems function in part as translators that mask incidental and accidental differences between languages in the access and retrieval process.

Computer intermediaries provide a set of services that usually go beyond translation. Frequently a richer and more consistent user environment adds customized value to the entire process. Uploading allows locally stored and maintained strategies to be sent to the host. Downloading allows results retrieved from host systems to be locally saved and processed. Assistance is given in selecting databases; online descriptions of the systems and databases are available; accounting subsystems are provided; and results can sometimes be transformed and factored back into queries.

Examples of mediating systems involving software running on standalone dial-up mainframes include the experimental CONIT, at MIT, and the former Chemical Substances Information Network (CSIN), funded by the National Library of Medicine. Another switching and mediating service that is dialed up but uses microprocessor hardware is EasyNet.[3]

Examples of software that can run on the user's own microcomputer include several packages that are no longer actively marketed: OL'SAM from the Franklin Institute, InSearch and ProSearch from the Menlo Corporation, and SearchMaster from SDC.[4] Microcomputer software currently available includes the Sci-Mate Searcher (Version 2.0) from the Institute for Scientific Information (ISI), Micro-CSIN and the Grateful Med from the National Library of Medicine, and Search Works from Online Research Systems.[5]

In developing ISI's Sci-Mate Searcher, the use and structure of network access, retrieval languages, and databases suggested distinct methods for automatic translation. The characteristics of use and structure and the methods developed to accommodate them will be described in the next four sections of this paper. Particular features of the Sci-Mate Searcher will not be described. Rather, general principles of automating online language translation will be described.

Two Interfaces: The User and the External System

When performing translation functions, an intermediate computer must manage two distinct language interfaces: one with the user and one with the external systems. Recognizing these two interfaces as distinct and isolating their distinctive operations in separate modules is essential for successful design.[6]

The user interface provides all significant retrieval functions and capabilities to the user. Controlled here is what the user can request and the way it can be requested. Also controlled here is what the user is given from the external system and the way it is presented.

"The external system" refers to everything beyond the serial port of the intermediate computer. The external system interface defines what modems, networks, and host systems are supported. It also defines the functions and capabilities that will be used in interacting with these external systems. The commands issued on the serial line are constructed in the external system interface modules, and the responses from the external systems are first processed by these modules too.

Intermediate translating software could be designed and written to directly convert user input into a form required by the current host and directly convert the current host responses into a standard form for the user. It is tempting to design and code this way. However, as more entities and capabilities are added to the external system, designing and writing yet another module for direct translation becomes quickly untenable; the mediation software soon becomes an intricate and incomprehensible network of code; and maintenance becomes impossible.

Design and implementation is much more easily managed if intermediate data structures are defined for all transactions. These structures store inputs from both users and systems in a standard form. The structures are accessed in separate steps for constructing acceptable commands for the network and host systems and for presenting consistently formatted information from the host to the user. The data structures serve as a buffer between the two interfaces. In the Sci-Mate Searcher, these data structures have been called the intermediate language.

With the intermediate language, user input destined for the host is accepted without regard to which host is currently online. Only in the system interface step that follows are the particulars for the current host added to data extracted from the intermediate structure. Similarly, responses from the current host are stored in the intermediate language data structures. These are then accessed in a separate step as host-independent data which are transformed and presented to the user.

Access Protocols: Description and Requirements

Access protocols here refers to the process of negotiating modems, networks, and host system passwords. Access is often viewed as a rote but necessary nuisance. It involves a series of steps which depend upon the details of particular hardware and systems. These steps include the following:

1. dial the network node, manually or through modem control;
2. inform the network about the terminal speed and type;
3. instruct the network about flow control and line padding;
4. specify the name or address of the host system;
5. negotiate the password(s) with the host; and
6. answer any standard host questions about news, etc.

Until the early 1980s, modems were usually dialed manually. Intelligent modems which allow software to control dialing now make it possible to fully automate this task. Among intelligent modems, models from D.C. Hayes have set the standard for control language. Data networks presently used in the United States for online access include Telenet, Tymnet, Uninet, Dialnet, and InfoNet.

The user's primary requirement for automated access is to get connected to a particular host or database. Users at all levels of expertise would like to log on by simply naming the database. All the steps of the process are amenable to complete automation. However, if a failure occurs and automatic access cannot be carried out, users should be able to choose alternatives such as trying again, quitting the session, or taking over manually.

A special-purpose procedural language was felt to be appropriate for automating access in the Sci-Mate Searcher. Access is algorithmic in that it has a clear beginning and end. Translation by a procedural language is effective in access where only a small number of possible messages from the external system can be anticipated at any given moment in the process. These messages, and the lack of any message in a given period of time, are known as potential states in the process. When a possible message is received, the state is said to be realized.

The actions to be performed when a potential state is realized are few and readily defined and programmed. These actions mainly involve a message in response. After specifying the host or database, the user becomes largely an observer during automated access. This allows the intermediate computer to control and respond to other computers, a relatively straightforward task in the case of access.

Retrieval Commands and Responses: Description and Requirements

The retrieval languages of the bibliographic host systems consist of commands entered by users and host responses to these commands. In the United States and Canada there are about half a dozen major host systems; in Europe there are at least six more. In each case, they accept a command consisting of a command verb (sometimes implied) followed by an argument.

"Retrieval command and response" will refer here to language components that control the host retrieval software but are independent of any specific database. Most retrieval systems provide at least the following basic commands:

1. pick a database or set of databases;
2. browse inverted indexes to the database(s);
3. select terms and specify term logic;
4. display records from sets constructed;
5. request records to be printed and mailed;
6. review the sets created during the session; and
7. leave the system.

Additional capabilities are frequently provided that build on these basic ones. They include commands to:

1. make selections from the inverted index display;
2. search for complete phrases in database records;
3. limit the search to years, updates, or languages;
4. specify ranges of records and formats for displaying and printing records; and

5. save strategies—recall and use saved strategies.

The retrieval software systems always allow a series of sets to be created. These sets consist of pointers to records which can be directly reviewed. The sets of pointers can also be used in the argument of the selection command. As terms in logical expressions they result in further sets as the search strategy is refined.

Automated translation of retrieval system commands must provide a consistent syntax in place of the broad diversity of construction required by the different systems accessed. This requires at the very least a single set of command verbs or function names to be used across systems. It further requires a unified, or at least consistent, set of conventions in the construction of command arguments. Finally, responses from the host should be standardized before being presented to the user.

Commands for the host systems are constructed from: (1) a standard command specified by the user, (2) data elements entered by the user, and (3) punctuation and other connecting elements required by the host. All of these are ordered as required by the host. Responses received from the host are parsed into the intermediate language tables from which significant data elements are extracted and reconstructed in a consistent form for presentation to the user.

The intermediate system should automatically enter and leave the "modes" found on hierarchically organized command systems, such as BRS and Questel. All information about sets created and commands issued in the current session are saved for the duration of the session. As part of response parsing, the intermediate system recognizes error messages and conditions and can assert failure when an excessive time delay occurs.

In developing Sci-Mate, directly coded routines have been found to be most effective for translating intermediate language data into multiple host system retrieval commands. Here, exceptions prevail over rules. Conversely, directly coded translations from the intermediate language into a unified user presentation have been found to provide more effective direct control than a meta-representation could possibly provide.

Retrieval languages come in families: Orbit and Elhill share a common origin and form one family; Recon, Dialog, and ESA-IRS/Quest form another; BRS and DataStar (Switzerland) form a third; and there are others.[7] In translating retrieval commands and responses, a matrix of functions by retrieval language family must be managed.

The only regular form or pattern in the commands across the language families is the verb-argument arrangement. Even here the verb is often implied especially in the selection function. The syntax of each argument in each language family follows no pattern that can be observed across language families. Figure 1 gives an example of just those cells of the matrix that contain the Dialog, BRS, and NLM commands for display. Figure 2 shows the various ways in which Dialog, BRS, and NLM report sets formed.

SYSTEM	VERB	ARGUMENT
Dialog	T or TYPE	<set #> / <format code> /<item list or range>
BRS	..P or ..PRINT	<set #> <format code> / DOC= <item list or range>
NLM	PRT or PRINT	SS <set#> <format> SKIP <first record # - 1>

Figure 1. Variations among host requirements for the Online Display Command.

SYSTEM	RESPONSE
Dialog	<set #> <postings count> <set formation expression>
BRS	RESULT <postings count> DOCUMENTS
NLM	SS <set #> PSTG <postings count>

Figure 2. Variations among host responses to Set Formation Commands.

Modularization remains an important principle and practice. In particular, the user interface in Sci-Mate provides a separate module for each function. The host interface contains routines for all host functions in a single module with each host function managed by one or more procedures. Both the host and user interfaces draw data from and provide data to the Intermediate Language data structures where standardization ultimately takes place.

Database Structure: Description and Requirements

Online bibliographic databases show many parallel structural characteristics even across hosts. The relative simplicity and consistency of their structures make it possible to define and store fairly complete information about the structure, but of course not about the content, of most databases.

The data records themselves are textual with variable length fields. All hosts have at least one format in which the fields are labeled with prefix tags of two to four characters. The labels are followed by one or more lines of textual data. The second and subsequent lines of data in a field are usually indented to show that they are part of the same field.

Inverted indexes provide retrieval keys for the data in most fields. A basic index containing terms from all fields is present on many hosts. Other fields have their own inverted indexes. Usually the contents of inverted indexes can be reviewed starting anywhere and continuing through the index in alphanumeric order.

There are three ways in which term coordination is handled in searching. Some fields can be searched using only single word terms. Other fields precoordinate two or more terms into searchable phrases in addition to single word terms. Finally, proximity or adjacency searching, in which terms are postcoordinated at the time the command is constructed, is allowed in most fields.

The user's primary requirement at the database level is information about the contents and structure of the database. Such information can guide the user in the selection of appropriate fields and the construction of search expressions. This information also specifies what tags to use to designate fields and the acceptable form for terms and expressions in each field.

This information can be found in the database provider's documentation and in the host system's fact sheets. After locating the information, the conventional searcher must switch attention between the manual and the terminal screen. Users can also experiment while online to determine the syntax allowed in a field, but this can be time-consuming and therefore costly. Conventional online searching can be enhanced by making these tools immediately available on the terminal screen.

A table of database information is used in the Sci-Mate Searcher to translate database syntax. Here the task consists of transforming one definition of the data structure into another. Both user and host requirements for information about the database are represented in a single entry in the table. The table may be called a "knowledge base" since it represents expert knowledge about databases.

In this knowledge base, the tags and usage of database elements required by the host are mapped to more complete names and descriptions for the same elements stored for users. Users face a mnemonic and encyclopedic problem in handling details of syntax at the database level. The knowledge base is intended to solve this problem with immediate information about the database.

In addition to information about the fields, certain global information about the database must also be stored. This includes: the name of the database as it is to be presented to the user; the name of the database as it is known to the host; and an indication of the host on which the database is found.

The knowledge base stores information for both the user and the host system. For the user, the complete names of the fields are presented as part of a menu selection. After a selection is made, stored descriptions of the fields and its subfields are given as prompts. The user is told whether or not phrases are allowed in the field.

For the host system, the value, placement, and punctuation of the field tag is extracted from the table and used in the construction of selection or browse commands. If phrases are allowed in the field, the appropriate adjacency or proximity symbols are supplied by the table.

Consistent User Presentation for the Retrieval Process

Three major language classes used in information retrieval have been identified: access protocols, retrieval commands and responses, and database structures. Numerous specific languages are found within each of these classes. For computer mediation, distinct translation methods were found to be most effective with each language class. How can these distinct translation methods be integrated for a consistent presentation to the user?

First, what is meant by a consistent user presentation? This at least means one in which the user is asked to become accustomed to a manageable, well-defined, and easily learned set of conventions. It also means one in which the transitions from one set of options to another fit logically and naturally into the user's experience and expectations. It does not mean that some particular device or method made possible by the hardware and software is necessarily employed.[8]

Over the past few years there has been a trend away from explicit and toward implicit language interfaces for users.[9] The oldest form of interactive interface and the one that traditional online information systems continue to take, is the host-prompt/user-command/host-response. This derives from the technology and economics of early timesharing. This form is one dimensional in that it looks like a simple dialogue alternating between a line from the user and one or more lines from the host system. Being mnemonic, it requires users to remember or quickly locate details about how the language can be used.

As smart terminals and microcomputers become available, much higher display transfer rates are possible. For user interaction, options can be economically presented on a menu. The two-dimensional surface of the video display unit (VDU), fully refreshed in less than a second, presents the user with explicit options for direct selection. Thus interaction requires less to be remembered and less to be entered, as a selection of an entry from a menu is sufficient. The machine becomes proactive as it supplies direct and verbal options to its user. Much software for microcomputers, including Sci-Mate, uses this method for interacting with users.

The newest class of interactive interfaces communicates in still more implicit language. New devices and methods give a spatial and environmental feeling to user interaction. Color graphics and icons communicate concepts without using words. Pop-up windows and pull-down menus present options instantly without obliterating the current context. Pointing devices such as mice allow metaphorical navigation to every location on the screen. Foreground/background activities make users feel that they have to wait less time for the machine to complete its tasks. Memory-resident utilities allow machines to be used for diverse tasks; applications are availale at the touch of a key.

No retrieval intermediary software developed to date has fully taken advantage of the new devices and techniques for interfacing with users. Almost all software in the mediation genre uses either a command- or menu-driven interface. However, consistency and ease of use are not precluded by either of these methods.

For access protocols, it is sufficient to provide users a method of selecting either the database or the host system. A further step is to automatically select a database or set of databases depending on the general subject area being researched. This was done by In-Search and is being done by EasyNet. Such a selection can be performed adequately by either a command-driven or menu-driven interface.

For retrieval commands and responses, users should be able to simply specify the function or operation to be performed. Better yet would be to provide users with recommendations for actions such as was done by the Individualized Instruction for Data Access (IIDA) project.[10] In either case, a menu of possible or recommended functions continually available to the user for selection is helpful.

Finally, for database structures represented by a knowledge base, a menu can effectively provide the field names for selection. The tables are further used to present prompts for both field and subfield data.

Conclusions

Three distinct classes of language have been identified in online bibliographic information retrieval. These classes are: access protocol, retrieval command/response, and database syntax. Computer intermediary systems that perform language translation should recognize the special problems and requirements posed by each class and adapt mediation to fit these problems.

In developing the Sci-Mate Searcher, it was found that a special-purpose procedural language was most effective for managing access protocol; direct computer routines drawing upon intermediate data structures were most effective in handling retrieval commands/responses, and a knowledge base was most effective in dealing with database syntax. Despite the different languages and methods for their translation, the user can and should be presented with a consistent view of the whole mediated process.

REFERENCES

1. Negus, A.E. "Development of Euronet Common Command-Language." *Online Review* 3(no. 4, 1979):414.

2. Nicholas, D., and Harman, J. "The End-User: An Assessment and Review of the Literature." *Social Science Information Studies* 5(no. 4, 1985):173-84.

3. Marcus, R.S., and Reintjes, J.F. "A Translating Computer Interface for End-User Operation of Heterogeneous Retrieval Systems. Part I. Design." *JASIS* 32(July 1981):287-303; Bracken, M.C. "Chemical Substances Information Network (CSIN) - Status." In *ASIS Proceedings*, edited by Roy D. Tally and Ronald R. Deultgen, p. 362. White Plains, N.Y.: Knowledge Industry Publications, 1979; and O'Leary, M. "EasyNet—Doing It All for the End-User." *Online* 9(July 1985):106-13.

4. Toliver, D.E. "OL'SAM: An Intelligent Front-End for Bibliographic Information Retrieval." *Information Technology and Libraries* 1(Dec. 1982):317-26; Quint, B. "Menlo Corporation's Pro Search—Review of a Software Search Aid." *Online* 10(Jan. 1986):17-25; and System Development Corporation. *The ORBIT SearchMaster System User Manual.* Santa Monica, Calif.: SDC, 1984.

5. Garfield, E. "The Integrated SciMate Software System. Part 1." *Current Contents* 28(Sept. 1985):3-10; National Library of Medicine. "Micro-CSIN Workstation" (Fact Sheet, Feb. 1986)(pamphlet); "Beta-test announcement, Grateful Med" (fact sheet developed by the National Library of Medicine)(pamphlet); and "SearchWorks Exciting Post-Search Packaging Search Aid Software." *Database End User* 2(Feb. 1986):32.

6. Tague, Jean. "The Two-Faced Interface." In *ASIS Proceedings*, pp. 81-87. White Plains, N.Y.: Knowledge Industry Publications, 1985.

7. Online. Inc., Pemberton, J.K. ed. *ONLINE International Command Chart* (spiral bound book). Weston, Conn.: Online, Inc., 1985 (chart).

8. Wallace, Danny P. "A Preliminary Examination of the Meaning of User Friendliness." In *ASIS Proceedings*, pp. 337-41. White Plains, N.Y.: Knowledge Industry Publications, 1985.

9. Newell, A. and S.K. Card. "The Prospect for Psychological Science in Human-Computer Interaction." *Human-Computer Interaction* 1(1985):209-42.

10. Meadow, Charles T., et al. "A Computer Intermediary for Interactive Database Searching. Part I. Design." *JASIS* 33(Sept. 1982):325-32.

LINDA C. SMITH

Associate Professor
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign

User Friendly Future: Applications of New Information Technology

What Is User Friendly?

This paper considers the clinic theme, "What Is User Friendly?" from a scientific and technical perspective. As Burch has observed in the introduction to a bibliography on computer ergonomics and user friendly design, the term *user friendly* is an anomaly as a technical term: "Most words borrowed from science enter the popular language stream long after their associated discoveries have become history. The term 'user friendly' is an exception to this rule; it became popular long *before* a scientific basis for 'user friendliness' had even been looked for."[1] The current emphasis on user friendliness is both market- and technology-driven. There is an interest in making computers more useful tools for people who are not computer specialists, thus expanding the potential user population; and there are new technological components that may be employed to make systems easier to use.

Definitions proposed for user friendly/friendliness range from brief dictionary definitions (e.g., "a system with which relatively untrained users can interact easily")[2] to lists of criteria (e.g., criteria for user friendliness proposed by Trenner and Buxton).[3] Although a review of these definitions and criteria is one means of providing a context within which to view new technological developments, this paper instead begins with a historical perspective, describing selected proposals for user friendly systems made over the past forty years.

Technology Forecasting: Techno-poetic Fantasies

In an essay introducing the technology section of *The New Encyclopaedia Britannica Propaedia,* Lord Ritchie-Calder remarks that: "From earliest time and beginning with the simplest contrivances, every discovery and invention has depended on the fact that the human being is not only a perceptual but also a conceptual creature capable of observing, memorizing, and juxtaposing images. He can make a mental design, a techno-poetic fantasy, even when the means of actually producing it are not available."[4] In the domain of information system design, there have been a number of such techno-poetic fantasies, designs for user friendly systems not realizable with the technology available at the time they were proposed. Rheingold has recently surveyed several of these proposals and the people behind them.[5] Those described briefly in the following paragraphs originated with Bush, Licklider, Engelbart, Nelson, and Kay: memex, procognitive systems, the augmented knowledge workshop, hypertext, and dynabook.

Vannevar Bush's article, "As We May Think," in which he proposed memex and other devices, has frequently been cited in the library and information science literature since it first appeared in *Atlantic Monthly* in July 1945.[6] Less well known is the condensed and illustrated version which appeared in *Life* 10 September 1945, including illustrations of future information technology such as memex.[7] Memex, as envisioned by Bush, is a mechanized private file and library. It is "a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."[8] Bush emphasized the value of organizing the contents using associative indexing, "whereby any item may be caused at will to select immediately and automatically another. This is the essential feature of the memex. The process of tying two items together is the important thing."[9]

In 1967 Bush had an opportunity to assess how much progress had been made toward the construction of memex.[10] He observed that: "Great progress...has been made in the last twenty years on all the elements necessary. Storage has been reduced in size, access has become more rapid. Transistors, video tape, television, high-speed electric circuits, have revolutionized the conditions under which we approach the problem."[11] However, Bush was not optimistic that a personal machine would be affordable in a short time. He did not foresee the rapid progress in integrated circuit technology which led to personal computers in the 1970s.

In 1965 J.C.R. Licklider published *Libraries of the Future* in which he described the likely characteristics of future computer-based information systems.[12] He coined the term *procognitive systems* to differentiate them from libraries, since the intent was that such systems "will extend farther into the process of generating, organizing, and using knowledge" through interaction among men, computers, and the body of knowledge.[13] Criteria to be met by procognitive systems include: converse or negotiate with the user while he formulates his requests and while responding to them; adjust itself to the level of sophistication of the individual user, providing terse streamlined modes for experienced users working in their fields of expertness, and functioning as a teaching machine to guide and improve the efforts of neophytes; provide the flexibility, legibility, and convenience of the printed page at input and output and, at the same time, the dynamic quality and immediate responsiveness of the oscilloscope screen and light pen.[14]

In 1982 Licklider had an opportunity to reflect on developments since 1965.[15] Although he noted considerable advances in the technological infrastructure, such as increased storage capacity and the availability of networks for digital transmission of information, he remarked that "the practically important application of information technology by libraries has not been, the past eighteen years, on any direct path to the procognitive system I was trying to describe in *Libraries of the Future*."[16] Nevertheless, he concludes by suggesting that, by the year 2000, librarians will have two important roles: (1) contributing to the work of the online intellectual community involved in generating and using the body of knowledge, and (2) organizing and maintaining the body of knowledge which will exist in electronic form.

In 1963 a series entitled "Vistas in Information Handling" began with a volume devoted to *The Augmentation of Man's Intellect by Machine*.[17] The lead paper in that volume, prepared by Douglas C. Engelbart, presented a conceptual framework for the augmentation of man's intellect.[18] At the recent Association for Computing Machinery (ACM) Conference on the History of Personal Workstations, Engelbart reviewed research conducted in the intervening years toward realizing the "augmented knowledge workshop"—the place in which a person finds the data and tools with which he does his knowledge work, and through which he collaborates with similarly equipped workers.[19] Engelbart feels that human knowledge work capability can be enhanced through properly harnessing this new technology. Although many of the technologies, both hardware and software, originally developed by Engelbart's group have now made their way into commercial products, he concluded his conference presentation on a somewhat pessimistic note: "I still don't see clear perceptions about what we humans can gain in new capabilities, or about how this may come about. There are constant, echoing statements about how fast and smart the computers are going to be, but not about how the enhanced computer capabilities will be harnessed into the daily thinking and working life of our creative knowledge workers."[20]

At a colloquium on information retrieval held in 1966, Theodor H. Nelson argued that access to information may not be best accomplished either by indexing techniques (document retrieval) or queriable information networks (content retrieval).[21] As an alternative, he suggested that digital text storage and display make possible the creation of hypertext or nonlinear text systems. Hypertext is the combination of natural language text with the computer's capacities for interactive, branching, or dynamic display; it "may differ from ordinary text in its sequencing (it may branch into trees and networks), its organization (it may have multiple levels of summary and detail), its mode of presentation (it may contain moving or manipulable illustrations, moving or flashing typography), and so on."[22] Nelson has been pursuing development of the technology required to support this concept, as reported in his book *Literary Machines.*[23]

The final techno-poetic fantasy noted here is the dynabook, proposed by researchers at the Xerox Palo Alto Research Center.[24] The dynabook would be "a personal dynamic medium the size of a notebook...which could be owned by everyone and could have the power to handle virtually all of its owner's information-related needs."[25] Alan Kay and Adele Goldberg describe what such a device would be:[26]

> Imagine having your own self-contained knowledge manipulator in a portable package the size and shape of an ordinary notebook. Suppose it had enough power to outrace your senses of sight and hearing, enough capacity to store for later retrieval thousands of page-equivalents of reference materials, poems, letters, recipes, records, drawings, animations, musical scores, waveforms, dynamic simulations, and anything else you would like to remember and change.

Although none of these authors used the term *user friendly* in characterizing the products of their imagination which are now at least partially realizable with available technology, a technologically based definition of the concept user friendly should include such visions of the future. In each case ease of interaction was taken as a given; instead the focus was on means of creating, organizing, searching, and using the contents of the knowledge base.

Technology Transfer: Information Technology

Before turning to a consideration of the technological components which will form the basis of user friendly systems in the future, it is appropriate to note the plethora of periodicals which have emerged in an effort to speed the transfer of technology into the library context. Titles include *Information Technology and Libraries, Program: News of Computers in Libraries, Small Computers in Libraries, Microcomputers in Information Management, Library Software Review, The Electronic Library, Electronic Publishing Review, Online, Online Review, Database, Library Hi Tech, Library Hi Tech News, Library Technology Reports, Information Retrieval and Library Automation, Advanced Technology/*

Libraries, and *Information Today.* Periodicals such as *Library Journal* and *Wilson Library Bulletin* also now have regular columns devoted to library uses of technology. Although sources in the computer science and engineering literature must be consulted to follow current research in information technology, possibilities for application are documented in a reasonably timely manner in the periodicals published for a library and information science audience. Given the rapidity with which new developments occur, the next section simply highlights some of the technological components currently available for design and construction of more user friendly systems.

Technological Components: Hardware and Software

Developments in hardware contribute to user friendliness by making many alternatives first feasible and then economical. Because users of most systems can be expected to be a heterogeneous group, choices in hardware allow alternative modes of access to be implemented for a given system. For example, microcomputers can be substituted for dumb terminals now that information processing technology has become relatively inexpensive. This enables the system to present alternative interfaces, such as one that is menu-driven rather than command-driven. Local processing also offers the possibility of implementing gateways to simplify access to multiple systems, masking differences which users may find hard to remember.

Telecommunications contributes to ease of interaction through the transmission speed which can be supported. New types of links using fiber optics can support higher speed and larger bandwidth so that more data can be transmitted at a faster rate. In addition there are now possibilities for integrating voice, text, image, and data communications.

New forms of storage media make possible local, self-contained information systems as an alternative to interactive access of remote databases. In particular the optical disks, such as CD-ROMs, offer large capacity storage for digital data as well as visual images. Because cost to use such systems is no longer a function of connect time to a remote computer, new types of interaction which would be too costly in systems charging for use by the minute are possible.

Input/output devices have the most direct impact on perceived user friendliness. Input is no longer confined to the QWERTY keyboard which anyone but the touch typist may find cumbersome to use. Touching (using touch screens) and pointing (using devices such as the mouse) can be used to indicate choices in menu-based systems. Output can use printers, plotters, and display screens with possibilities for different fonts, colors, windows, and graphics. Although not yet as common, limited voice input and speech output allow the use of sound rather than tactile and visual means of recording and reporting.

Software is of course required to make all these hardware components operate. In judging user friendliness, one is concerned with what Shackel has termed the "cognitive and software interface."[27] Components include languages (e.g., use of command languages *v.* natural language), information organization, display format and layout, dialogue structure and design, error message design, and advanced interfaces (e.g., intelligent systems adaptive to the user). Tools are beginning to be available with which to design and build many of these components as identified, for example, in Bundy's *Catalogue of Artificial Intelligence Tools.*[28]

Given this wide range of technological components, the challenge is to combine elements to create more user friendly systems. As Smith notes, there are significant differences between designing hardware and software for the user interface.[29] Formal standards may be applicable to hardware design, but flexible design guidelines rather than standards are applicable to software design. For example, Rubinstein and Hersh present a well-developed set of guidelines for human-oriented design.[30] In general, more guideline information is available relating to the physical interface than to the cognitive interface.[31]

Technological Integration: Personal Workstations

Development of personal workstations represents the computing environment which will form the basis for user friendly systems in the future. The transition has been characterized by Perlis and White: "Twenty five years ago computing was stationary, ponderous and centralized. Its dominant role was to serve the critical needs and purposes of organizations and the sciences. Today matters are very different. Computation is personal, ubiquitous and expansive. Power is being supplied at and to the fingertips of the individual."[32] The workstation concept is sustained by four technologies: dedicated microprocessors, local area networks, local databases, and gateways to mainframes.[33] Various input/output devices are provided, depending on the tasks which the workstation is designed to support. The workstation is used to carry out both generic activities (e.g., calculation, word processing, mail) and profession-related activities (e.g., scientific or engineering analyses) with appropriate software support.

These computing and communication systems are already appearing in organizations of which libraries are a part, such as universities. At Carnegie-Mellon University, for example, a system named ANDREW is being developed with personal computers, raster graphics, high bandwidth communications, and time-sharing file systems as components.[34] The designers anticipate that ANDREW will affect university education in four main areas: computer-assisted instruction, creation and use of new tools, communication, and information access. With respect to information access, the designers comment that "a mark of tomorrow's professional will be the ability to navigate in large information repositories" including the library's database, worldwide databases, and databases developed within the university.[35] Some predictions of how such systems will be used have already appeared. For example, Spinrad offers what he terms "vignettes" describing how a typical student, professor, and administrator would function in an electronic university,[36] and Lancaster describes how the scientist could use an electronic information system to create, transmit, and receive information.[37] Some of the "techno-poetic fantasies" cited earlier also suggest ways in which a personal workstation could be used.

Technology Assessment: An Appropriate Skepticism

To provide a balanced discussion of technology in support of user friendliness, it is necessary to interject what John Shelton Lawrence has termed "appropriate skepticism." In discussing the use of computers for word processing, he notes that: "Computer users often allow their exhilaration with hardware and productivity to displace the critical attention they formerly gave to their manually produced material....The physical appearance of the computer's output is seductive in this regard; because it prints absurdity as beautifully as the most carefully wrought expression, one is tempted not to look beneath its surface."[38] A similar danger exists in the context of user friendly catalogs and other information systems. Problems may arise if the following factors are not taken into consideration.

Comprehensibility. In a piece entitled "Black Box Blues," Dixon remarked that "the real danger of the microelectronic era is posed by what was called, even in the days of macroelectronics, the black box mentality: passive acceptance of the idea that more and more areas of life will be taken over by little black boxes whose mysterious workings are beyond our comprehension."[39] The algorithms followed by computers are not necessarily comprehensible to users. Yet by knowing the basis for system decisions, the user can more appropriately accept, reject, or modify them. Designers must determine the extent to which computer processes should be made explicit rather than hidden.

Scope of the system. A great deal of effort can be expended to no purpose if the user seeks information which in fact is not contained in the system. In order to use the system intelligently, a user needs to understand its scope—i.e., the broad class of questions to which the system is designed to respond.

Limitations of the system. The attempt to make human-computer dialogues more like human-human dialogues may lead to an overly anthropomorphic interpretation of the computer system by users. Without a way to probe the limits of capabilities of a human-like system, the user is likely to attribute more power to it than it actually has.

Source of information. When information is sought from printed sources or from other people, the inquirer has some basis for judging the authoritativeness of the material or the response. By masking aspects of the search process from the user—such as database selection—and by presenting isolated responses—whether citations or facts—the inquirer has no basis for judging the domain covered or the reliability of the response.

Mastery of the system. In a piece entitled "Can Online Catalogs Be Too Easy?" Arret points out that user easy is not user friendly if progressive learning and system mastery are sacrificed.[40] If there is no way for the user to advance beyond the simple searches supported by the user friendly interface, then there is no way that the full power of the system can be exploited.

In the spirit of technology assessment, a discussion of the technology supporting user friendly systems must acknowledge these potential problems. Given the current limitations of user friendly systems, users must develop an appropriate skepticism and designers must explore approaches to deal with issues such as those enumerated earlier.

Halfway Technology Versus High Technology

In an essay on the technology of medicine written in 1971, Thomas introduced a distinction between what he termed "halfway technology" and "high technology."[41] He explained that halfway technology is characterized by things done after the fact in efforts to compensate for the incapacitating effects of certain diseases. He noted that the real high technology of medicine comes as the result of a genuine understanding of disease mechanisms, allowing prevention and/or effective treatment. Interpreting these concepts in the context of information technology, one could describe efforts to design more user friendly interfaces to existing systems as halfway technology, trying to improve access to systems not initially designed from the perspective of user needs. To achieve high technology, research is required to understand the needs of the user far better than is the case today. This theme is echoed by Chapanis who talks of "taming and civilizing computers" by discovering enough about human behavior to design computer systems for enhancement and enrichment,[42] and by Birnbaum who notes that the "domestication of microelectronics" will only be achieved by developing computer technology in the context of what the user wants to do.[43] At present the hardware is far ahead of theory and research in user customization. Fortunately, there is an increasing amount of interest and research activity in this area, drawing on behavioral scientists as well as computer scientists.

User Friendly Future

This discussion began with the observation that user friendly is an anomaly as a technical term. Nickerson has suggested a simple alternative which may prove more satisfying:[44]

> Whether "friendliness" is the right concept is perhaps a matter of taste. "Usability" strikes me as the more appropriate and completely adequate concept; in imputing the quality of friendliness to a machine, one is diluting the meaning of one of the most pleasant of words.

And Burch in turn offers a measure of usability:[45]

> *System transparency* is the ultimate, ideal measure of computer usability. It is achieved when a system's overall design is so compatible with the way the user thinks, talks, listens, remembers, perceives, processes information, asks questions, makes decisions, and solves problems, that the system itself requires none of the user's attention and, in effect, becomes invisible. It happens in the same way that a reader curled up with a good book becomes unaware of the paper, the typeface, the book itself, or the room around him.

The current concern for user friendliness can be viewed as an attempt to cope with halfway technology. Future attention to usability and usefulness may lead the way toward high technology.

ACKNOWLEDGMENT

The author is indebted to David N. King, who introduced her to Thomas's discussion of halfway and high technology.

REFERENCES

1. Burch, John L. *Computers: The Non-Technological (Human) Factors: A Recommended Reading List on Computer Ergonomics and User Friendly Design*. Lawrence, Kansas: The Report Store, 1984, p. 6.

2. Meadows, A.J., et al. *Dictionary of Computing and New Information Technology*. London: Kogan Page, 1984, p. 211.

3. Trenner, L., and Buxton, A.B. "Criteria for User-Friendliness." In *9th International Online Information Meeting Proceedings*, pp. 279-87. Oxford: Learned Information, 1985.

4. Ritchie-Calder, Lord. "Knowing How and Knowing Why." In *The New Encyclopaedia Britannica Propaedia*, 15th ed. Chicago: Encyclopaedia Britannica, 1986, p. 261.

5. Rheingold, Howard. *Tools for Thought: The People and Ideas Behind the Next Computer Revolution*. New York: Simon & Schuster, 1985.

6. Bush, Vannevar. "As We May Think." *Atlantic Monthly* 176(July 1945):101-08.

7. ________. "As We May Think: A Top U.S. Scientist Foresees a Possible Future World in Which Man-Made Machines Will Start to Think." *Life* 19(10 Sept. 1945):112-14, 116, 118, 121, 123-24.

8. ________, "As We May Think," pp. 106-07.

9. Ibid., p. 107.

10. ________ . "Memex Revisited." In *Science Is Not Enough*, pp. 75-101. New York: William Morrow & Company, 1967.

11. Ibid., pp. 99-100.

12. Licklider, J.C.R. *Libraries of the Future*. Cambridge: MIT Press, 1965.

13. Ibid., p. 6.

14. Ibid., pp. 36-37.

15. ________ . "The View from the Half-Way Point on a Journey to the Future. A Progress Report on the Interaction between Libraries and Information Technology." In *Large Libraries and New Technological Developments*, edited by C. Reedijk et al., pp. 13-34. München: K.G. Saur, 1984.

16. Ibid., p. 16.

17. Howerton, Paul W., and Weeks, David C., eds. *The Augmentation of Man's Intellect by Machine* (Vistas in Information Handling, vol. 1). Washington, D.C.: Spartan Books, 1963.

18. Engelbart, Douglas C. "A Conceptual Framework for the Augmentation of Man's Intellect." In *The Augmentation of Man's Intellect by Machine*, pp. 1-29.

19. ________ . "The Augmented Knowledge Workshop." In *ACM Conference on the History of Personal Workstations*, edited by John R. White, pp. 73-83. New York: Association for Computing Machinery, 1986.

20. Ibid., p. 81.

21. Nelson, Theodor H. "Getting It Out of Our System." In *Information Retrieval: A Critical View*, edited by George Schecter, pp. 191-210. Washington, D.C.: Thompson Book Company, 1967.

22. Ibid., p. 195.

23. Nelson, Ted. *Literary Machines*. Swarthmore, Pa.: Ted Nelson, 1981.

24. Kay, Alan, and Goldberg, Adele. "Personal Dynamic Media." *Computer* 10(March 1977):31-41.

25. Ibid., p. 31.

26. Ibid.

27. Shackel, B. "Ergonomics in Information Technology in Europe—A Review." *Behaviour and Information Technology* 4(Oct.-Dec. 1985):263-87.

28. Bundy, Alan, ed. *Catalogue of Artificial Intelligence Tools*, 2d ed. Berlin: Springer-Verlag, 1986.

29. Smith, Sidney L. "Standards versus Guidelines for Designing User Interface Software." *Behaviour and Information Technology* 5(Jan./March 1986):47-61.

30. Rubinstein, Richard, and Hersh, Harry. *The Human Factor: Designing Computer Systems for People*. Burlington, Mass.: Digital Press, 1984.

31. Nickerson, Raymond S. *Using Computers: The Human Factors of Information Systems*. Cambridge: MIT Press, 1986, p. 226.

32. Perlis, Alan, and White, John R. "Foreword." In *ACM Conference on the History of Personal Workstations*, p. v.

33. Chorafas, Dimitris N. *Personal Computers and Data Communications*. Rockville, Md.: Computer Science Press, 1986, p. 12.

34. Morris, James H., et al. "Andrew: A Distributed Personal Computing Environment." *Communications of the ACM* 29(March 1986):184-201.

35. Ibid., p. 184.

36. Spinrad, Robert. "The Electronic University." In *Cohabiting with Computers*, edited by Joseph F. Traub, pp. 43-57. Los Altos, Calif.: William Kaufmann, 1985.

37. Lancaster, F. Wilfrid. *Toward Paperless Information Systems*. New York: Academic Press, 1978.

38. Lawrence, John Shelton. *The Electronic Scholar: A Guide to Academic Microcomputing*. Norwood, N.J.: Ablex, 1984, pp. 166-67.

39. Dixon, Bernard. "Black Box Blues." *The Sciences* 24(March/April 1984):11-12.

40. Arret, Linda. "Can Online Catalogs Be Too Easy?" *American Libraries* 16(Feb. 1985):118-20.

41. Thomas, Lewis. "Notes of a Biology-Watcher: The Technology of Medicine." *New England Journal of Medicine* 285(9 Dec. 1971):1366-68.
42. Chapanis, Alphonse. "Taming and Civilizing Computers." *Annals of the New York Academy of Sciences* 426(1 Nov. 1984):202-19.
43. Birnbaum, Joel S. "Toward the Domestication of Microelectronics." *Communications of the ACM* 28(Nov. 1985):1225-35.
44. Nickerson, *Using Computers*, p. 152.
45. Burch, *Computers*, p. 17.

INDEX